CHARACTERIZATION OF TTN SPLICING VARIANTS TO GAIN FURTHER INSIGHTS INTO THEIR ROLE IN SKELETAL MUSCLE DISORDERS

TALHA QURESHI

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

BIO- JA YMPÄRISTÖTIETEELLINEN TIEDEKUNTA
BIO- OCH MILJÖVETENSKAPLIGA FAKULTETEN
FACULTY OF BIOLOGICAL AND ENVIRONMENTAL SCIENCES

Degree Program in Molecular Biosciences
Master's Thesis
Helsinki
August 2019

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta – Fakultet – Faculty Faculty of Biological and Environmental Sciences		Koulutusohjelma – Utbildingsprogram – Degree Programme Master´s program in Molecular Biosciences
Tekijä – Författare – Author Talha Qureshi		
Työn nimi – Arbetets titel – Title Characterization of *TTN* splicing variants to gain further insights into their role in skeletal muscle disorders		
Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track Human Genetics		
Työn laji – Arbetets art – Level Masters	Aika – Datum – Month and year August 2019	Sivumäärä – Sidoantal – Number of pages 62

Tiivistelmä – Referat – Abstract

The TTN gene encodes a giant muscle protein called titin that regulates the function of muscle sarcomere and interacts with several other muscle proteins. Mutations in TTN are associated with a broad range of skeletal and cardiac muscle disorders termed titinopathies.

Previous studies have shown the importance of unusual TTN splicing events in patients with TTN-related cardiomyopathies and muscular dystrophies. In this project, we characterized eight TTN splicing variants to further expound on the pathogenesis of titinopathies and to enhance the diagnostic accuracy for patients with TTN mutations. In addition, we also made a comparative analysis of five different RNA/cDNA sequencing techniques to extrapolate on which approach is most suitable to study splicing variants in TTN gene.

Skeletal muscle samples of six patients were analyzed in this study who were previously detected with TTN variants in a compound heterozygous state from a targeted next-generation sequencing assay. Our results from traditional Sanger sequencing methods, second-generation (Illumina RNA-Sequencing) and third-generation sequencing (Single-molecule real-time sequencing) methods showed distinct splicing events in the form of partial or complete exon skipping, intron retention, and in few instances showed multiple splicing effects rendered by a single variant.

Complying with the guidelines of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology, the splicing variants were classified as pathogenic, likely pathogenic or variant of uncertain significance primarily on the basis of our experimental data.

To address which sequencing method is most promising for analyzing TTN splicing variants, Illumina RNA Sequencing is very efficient, though, the combination of Illumina RNA Sequencing with long-read sequencing could be ideal.

Our results further demonstrate that a near full-length titin is vital for survival until birth, and further studies are needed to understand the pathophysiology mechanism of congenital titinopathies.

Avainsanat – Nyckelord – Keywords
Alternative splicing, Titinopathies, *TTN,*

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Marco Savarese and Per-Harald Jonson

Säilytyspaikka – Förvaringställe – Where deposited
E-Thesis, University of Helsinki

Muita tietoja – Övriga uppgifter – Additional information

GRAPHICAL ABSTRACT

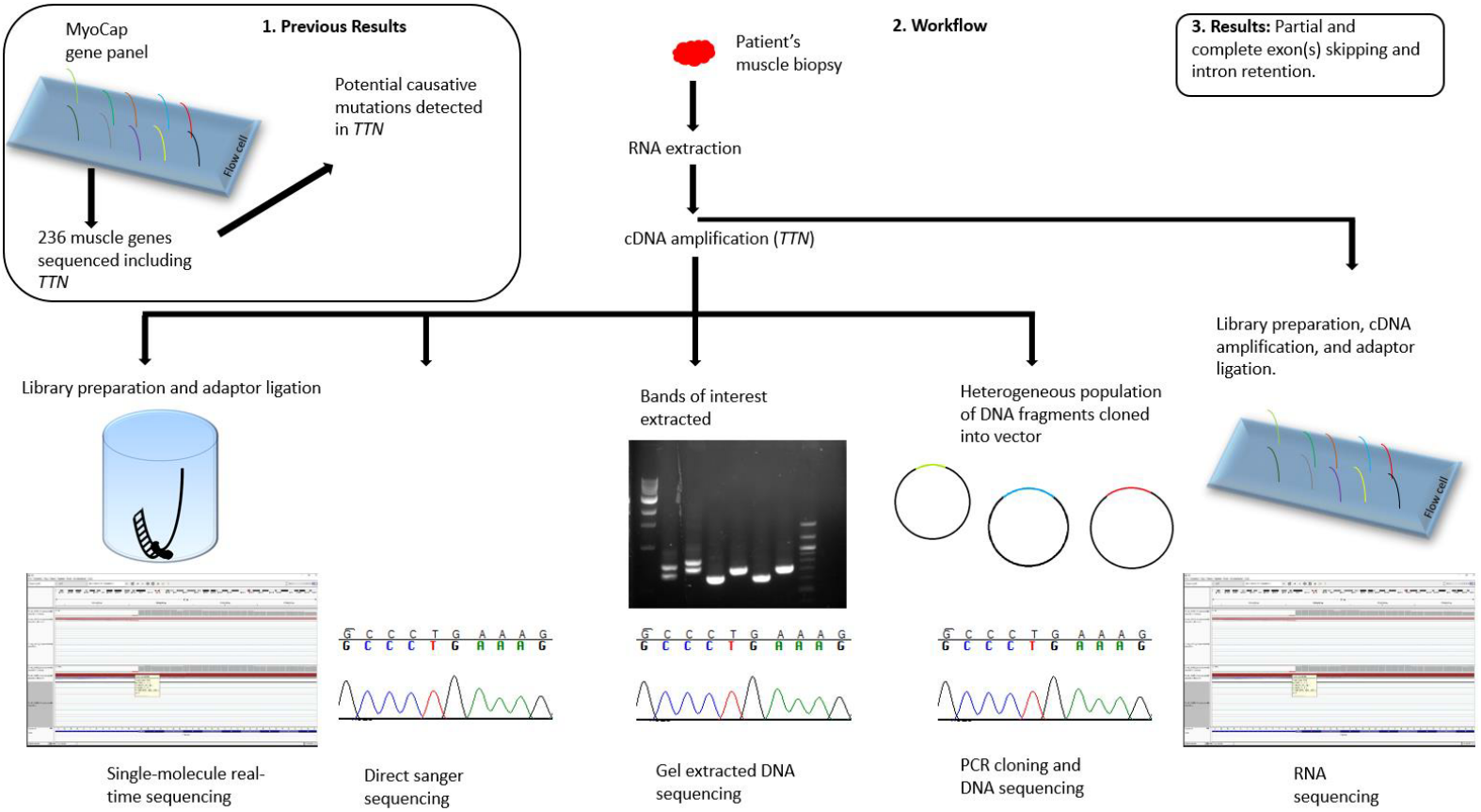

TABLE OF CONTENTS

Figures

Tables

GLOSSARY OF ABBREVIATIONS

ACMG/AMP	American College of Medical Genetics and Genomics and Association for Molecular Pathology
AS	Alternative splicing
ASEs	Alternative splicing events
cDNA	Complementary DNA
C-terminal	Carboxyl-terminal
DCM	Dilated cardiomyopathy
FN3	Fibronectin type III domain
gDNA	Genomic DNA
HMERF	Hereditary myopathy with early respiratory failure
hnRNPs	Heterogeneous nuclear ribonucleoproteins
Ig	Immunoglobulin-like
IGV	Integrative Genomics Viewer
kb	Kilobase
LB	Lysogeny broth
LOF	Loss of function
Mdm	Muscular dystrophy with myositis
MARPs	Muscle ankyrin repeat proteins
Murf2	Muscle-specific RING finger protein 2
NCBI	National center for biotechnology information
NGS	Next-generation sequencing
NMD	Non-Sense Mediated Decay
N-terminal	Amino-terminal
PEVK	Proline, glutamic acid, valine, lysine
PKG	Protein kinase G
Poly (A)	Polyadenylation
Poly (E)	Glutamic acid chain

PPAK	Proline, proline, alanine, lysine
PSI	Percent/proportion spliced in
PTC	Premature termination codon
RBM20	RNA-binding motif protein 20
RNA-Seq	RNA sequencing
RPM	Revolutions per minute
RT	Reverse transcription
SH3 domain	Src homology-3 domain
SMRT	Single molecule real time sequencing
SNP	Single nucleotide polymorphism
snRNA	Small nuclear RNA
SNVs	Single nucleotide variants
SRF	Serum response factor
S100A1	S100 calcium-binding protein A1
TK	Titin kinase
TTN	Titin gene
TTNtv	Titin truncating variant
UTR	Untranslated region
VUS	Variant of uncertain significance
WT	Wild type

1 INTRODUCTION

The TTN gene (cytogenetic location 2q31.2) encodes a giant muscle protein called titin that extends from Z-disc to M-band within the sarcomere and functions as a molecular spring during sarcomere contraction, relaxation and extension (Linke 2000) (Figure 1a and 1b). With a colossal size of approximately 300 kilobase (kb) pairs and 363 coding exons (Bang, Centner et al. 2001) (by far the most exons in any gene), TTN codes for the largest known protein to date and via alternative splicing can theoretically produce more than a million splice variants (Guo, Bharmal et al. 2010). However, only seven main human TTN isoforms have been reported according to the NCBI RefSeq transcript/nucleotide database.

Titin is primarily expressed in skeletal and cardiac muscles, but its expression has been reported in non-muscle and smooth muscle cells with some distinctive structural and interactive features than those of skeletal and cardiac titins (Keller, Eilertsen et al. 2000). The variability among titin isoforms is considerably high, and tissue-specific isoforms also exist (Guo, J Bharmal et al. 2010). Moreover, RNA sequences from cardiac samples have unraveled the presence of different isoforms within the same tissue (Roberts, Ware et al. 2015). All titin isoforms are approximately 27000 – 36000 amino acid residues long except the Novex 3 isoform that is merely composed of 5604 amino acid residues (Linke 2018). Clinically, a broad range of neuromuscular disorders, including skeletal and cardiac muscle disorders are ascribed to defects in TTN gene (Savarese, Sarparanta et al. 2016).

Neuromuscular disorders encompass a broad spectrum of genetically heterogeneous conditions that affect muscles. Gene table of neuromuscular disorders (musclegenetable.fr) reports 535 different genes associated with neuromuscular disorders out of which approximately 100 genes are mutated in primary myopathies. In routine diagnostics, traditional sequencing methods are not adequate to analyze patients with less characterized phenotypes that require sequencing of a vast number of genes including some very large genes such as TTN, NEB (codes for nebulin), etc in order to make an accurate diagnosis. However, this obstacle has been overcome with the advent of Next-Generation Sequencing (NGS), which allows hundreds of genes to be sequenced in a single batch (Schrijver, Aziz et al. 2012). Previously our group had developed a targeted NGS panel called

MyoCap, which contained exon and untranslated region (UTR) sequences of 180 genes pertinent to different myopathies (Evila, Arumilli et al. 2016). Later versions of MyoCap are upgraded with exon sequences of around 300 myopathy and muscular dystrophy-related genes, including UTR and intronic sequences of certain genes.

In spite of powerful sequencing technologies, the consequences of splicing mutations are better comprehended with transcriptome studies. In this project, we applied RNA and different cDNA sequencing techniques for characterization of specific splicing mutations detected in NGS MyoCap assay to discern whether a particular splicing variant/variant of uncertain significance (VUS) is benign or causative. The other variants in trans were not further included in this study as they were already predicted to be pathogenic according to the guidelines of American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP).

2 LITERATURE REVIEW

2.1 THE STRUCTURE AND FUNCTION OF TITIN

A single titin molecule is more than 1 μm in length and covers half of a sarcomere and regulates its stability during muscle extension and contraction (Tskhovrebova and Trinick 2010). The titin molecules run antiparallel to each other in a sarcomere, and the last few amino acids of N- termini of titins overlap in adjacent sarcomeres and within a sarcomere titin polypeptides overlap in their C-termini (Figure 1a), hence forming a consecutive filament array in myofibrils (McElhinny, Kakinuma et al. 2002). Being the third most abundant muscle protein after myosin and actin (Kellermayer, Smith et al. 2017), titin also serves as a molecular hub for signal transduction and a scaffold protein, harboring several binding sites for other proteins and aiding in myofibrillar assembly (Krüger and Kötter 2016).

The structure of titin is highly modular and comprised of four parts, categorized on the basis of their sarcomeric location, the amino-terminal (Z-disc anchored region), the I-band and A-band regions and the carboxyl-terminal (M-band anchored region) (Meyer and Wright 2013).

Approximately 90% of titin is composed of Immunoglobulin-like domains (Ig domains) and Fibronectin type III domains (FN3 domains) that are arranged predominantly in tandem repeats (Toro, Olivé et al. 2013), the other 10% is shaped by Z repeats, a serine-threonine kinase domain, PEVK domain and unique sequences.

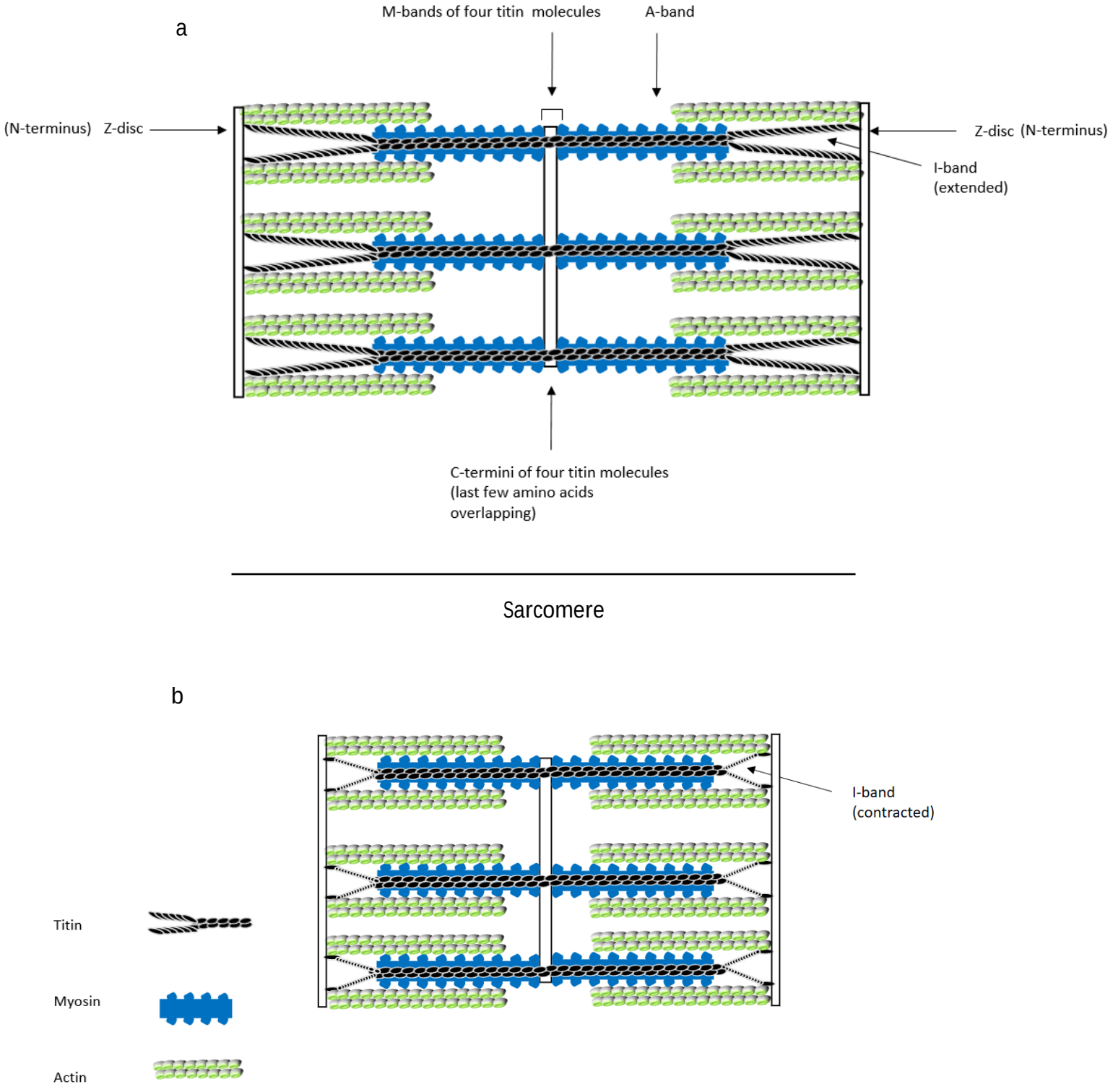

Figure 1: (a) The antiparallel layout of titin filaments within a sarcomere, a single titin polypeptide spans half of a sarcomere. (b) General depiction of how the elastic property of sarcomere is regulated by titin.

2.1.1 TITIN Z-DISC

The Z-disc region of titin is situated in the N-terminal of the molecule and encoded by the first few TTN exons. Two important Ig domains (Z1 and Z2) are positioned on the extreme N-terminal end of the molecule where these domains from two individual titin molecules interact with telethonin protein. The Z1 and Z2 domains are followed by unique sequence and Z repeat region (Figure 2). The Z repeat is a 45 amino acids multiple motif (encoded by TTN exons 8-14) that binds with the C-terminal region of α-Actinin, which crosslinks the actin filament to titin (Young, Ferguson et al. 1998, Linke 2008). The titin-telethonin interaction is highly resistant to mechanical forces during sarcomere extension and together with Alpha-actinin plays a key role in providing structural integrity to the Z-disc of sarcomere (Linke 2008).

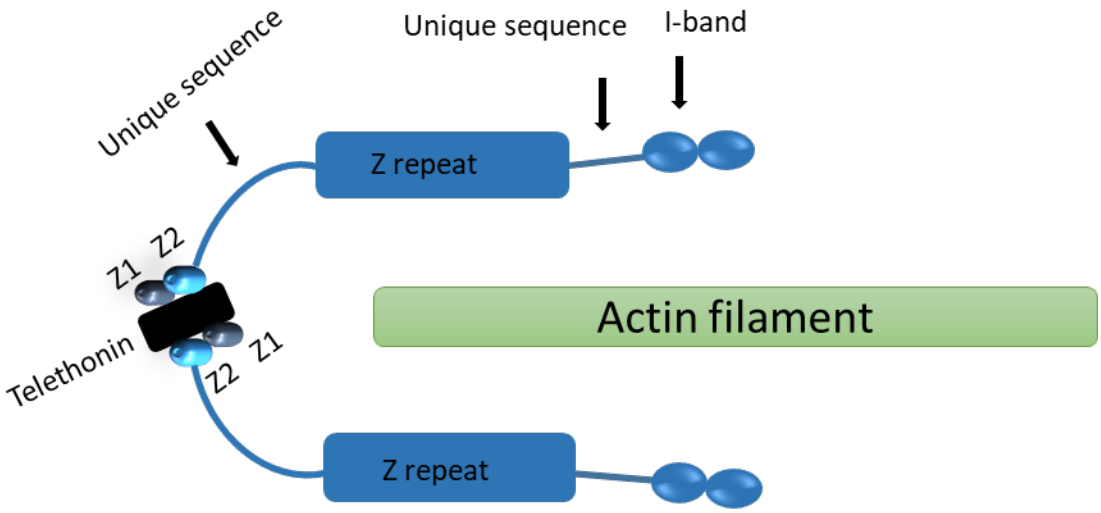

Figure 2: The figure illustrates the interaction of N-termini of two titin molecules with telethonin from the same half sarcomere.

Redrawn from http://www.ks.uiuc.edu/Research/telethonin/

2.1.2 TITIN I-BAND

The I-band region of titin (encoded by TTN exons 28-252, skipping exon 48 which is exclusive to Novex 3 isoform only) is largely composed of proximal (near Z-disc) and distal (near A-band) tandem Ig domains flanking a PEVK (Proline, Glutamic acid, Valine, Lysine) rich fragment with unique sequences interspersed (Figure 3). Large macromolecular conformational changes take

place in I-band upon relaxation and stretch (Meyer and Wright 2013). Additionally, force-induced phosphorylation of titin at specific sites in I-band modifies the myofilament stiffness (Hamdani, Herwig et al. 2017).

The PEVK domain, which is an epitome of molecular elastic, contains repeating motifs of 26-28 residues, the positively charged PPAK motifs (Lysine abundant motifs) and the negatively charged Poly (E) motifs (Glutamic acid abundant motifs). The opposite charges of motifs enable to overcome repulsive forces and are hypothesized to be crucial for ionic interactions between two classes of motifs (Greaser 2001).

The calcium-dependent endopeptidase calpain 1 binds to the proximal Ig domain of I-band (Raynaud, Fernandez et al. 2005) while actin, which is one of the most abundant muscle proteins is known to interact with PEVK segment and contributes to passive stiffness (Kulke, Fujita-Becker et al. 2001). Another giant muscle protein nebulin interacts with titin PEVK segment through its SH3 domain (Ma and Wang 2002); however, the exact function of titin-nebulin interaction is obscure at present (Chu, Gregorio et al. 2016). The Ig and PEVK domain in I-band collectively define the elasticity property of titin (Trombitas, Greaser et al. 1998).

The I-band also contains two unique elements called N2A (4 Ig domains and 106 amino acids unique sequence) and N2B (3 Ig domains and 572 amino acids unique sequence) (Cazorla, Freiburg et al. 2000). The N2A region is expressed in both skeletal and cardiac muscles, while N2B is only expressed in cardiac muscles (Labeit and Kolmerer 1995). The N2A region possesses binding sites for muscle ankyrin repeat proteins (MARPs) that have a role in muscle stress-activated pathway (Miller, Bang et al. 2003). Deletion of specific exons of N2A region disrupts a putative binding site for calpain 3, which leads to muscular dystrophy with myositis (mdm) in mice (Garvey, Rajan et al. 2002).

Likewise, the N2B region contains a protein-protein interaction site for transcriptional co-regulator called Four and Half LIM Domain 2 (FHL2), which might have role in directing metabolic enzymes to high energy demand sites in cardiac sarcomere (Lange, Auerbach et al. 2002). A stress protein AlphaB-crystallin (chaperone in nature) also interacts with N2B element. Studies in mouse have shown that deletion of N2B element prompts diastolic dysfunction and cardiac atrophy (Radke, Peng et al. 2007).

2.1.3 TITIN A-BAND

The A-band is the largest region of titin and is encoded by TTN exons 253-358. The structural organization of A-band is highly repetitive of Ig-like domain and FN3 domain (Figure 3). Both of these domain classes are arranged in 7 (D-zone) and 11 (C-zone) super repeats array, of which the C-zone is more frequent with a pattern of Ig-Fn3-Fn3-Ig-Fn3-Fn3-Fn3-Ig-Fn3-Fn3-Fn3 repeats. The A-band segment interacts with Myosin and Myosin binding protein-C, this ternary complex is believed to be crucial for sarcomere thick-filament assembly (Freiburg and Gautel 1996, Tonino, Kiss et al. 2017).

Contrary to I-band, the A-band region is relatively inextensible (Tskhovrebova and Trinick 2003). The A-band fragment is under the influence of active and passive forces during sarcomere extension and contraction, and the mechanical stress may subsequently trigger the titin signaling events that modulate muscle hypertrophy (Lewartowski and Mackiewicz 2014).

Since A-band titin serves as a platform for interaction of sarcomere thick filament proteins, it is characterized as a molecular ruler that governs the length of myosin filaments (Linke 2008).

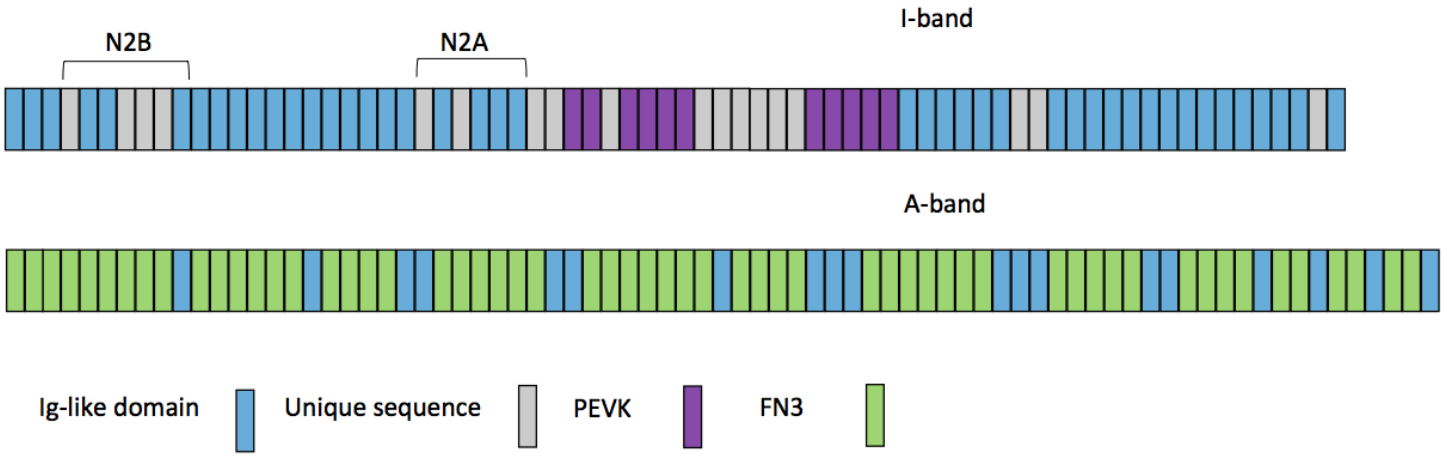

Figure 3: General overview of different domains of titin in I-band and A-band. The N2A region (expressed in both cardiac and skeletal muscles) and cardiac muscle specific N2B region are also shown.

2.1.4 TITIN M-BAND

The M-band region of titin is encoded by the last 6 TTN exons (359-364/Mex1-Mex6) and positioned in the C-terminal end of the molecule. In addition to Ig-like and FN3 domains, the M-line titin includes an autoinhibitory serine/threonine kinase domain (Labeit, Gautel et al. 1992), which inhibits its activation by C-terminal regulatory tail in the catalytic subdomain that intercepts the ATP binding; thus, titin kinase (TK) remains inactivated in a relaxed sarcomere. The mechanical force transmission removes the steric blockage of ATP binding site followed by autophosphorylation of a tyrosine residue that ultimately activates TK (Puchner, Alexandrovich et al. 2008) (Figure 4).

TK in its activated state interacts with numerous proteins including MURF2 and couples a signalosome. In the absence of mechanical stress, the signalosome is disengaged, and MURF2 is translocated to the nucleus where it can interact with a transcription factor called serum response factor (SRF) and hamper its transcriptional regulation activity. Hence, force-induced TK activation keeps the MURF2 out of the nucleus and augments the SRF-dependent muscle gene expression (Gautel 2011).

Titin M-band is also essential for sarcomere structural integrity and myofibril formation. Homozygous deletion of titin M-band results in failure of cardiac progenitor cells to differentiate and precludes sarcomerogenesis (Musa, Meek et al. 2006). Besides its significance in signaling and muscle development, mutations in titin M-band are linked to various neuromuscular disorders (Udd, Vihola et al. 2005, Carmignac, Salih et al. 2007, Chauveau, Bonnemann et al. 2014).

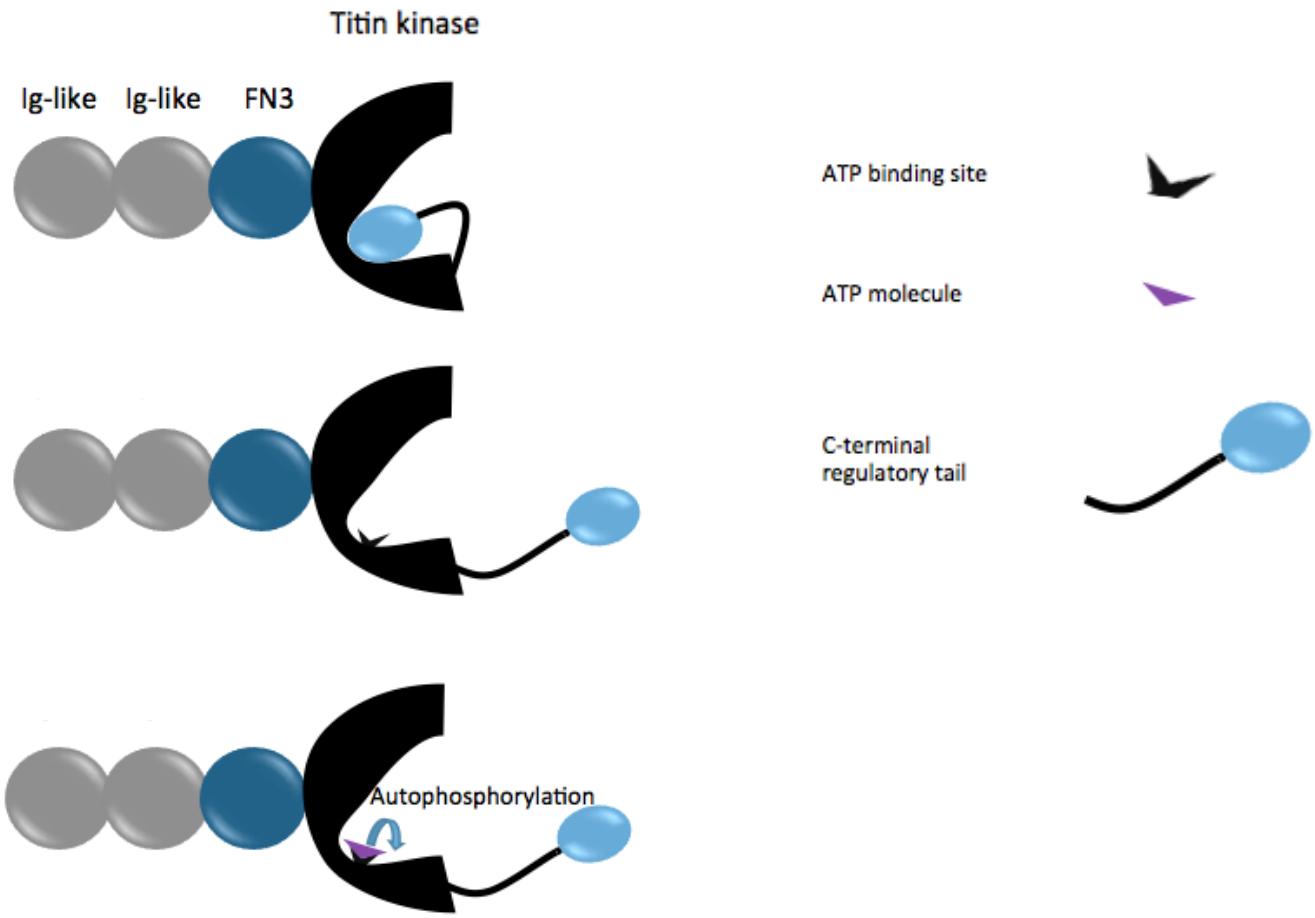

Figure 4: The figure encapsulates stepwise activation of titin kinase domain via autophosphorylation induced by mechanical force. Step 1: The ATP binding site is blocked by C-terminal regulatory tail in a relaxed sarcomere. Step 2: The mechanical force disengages C-terminal regulatory tail, exposing the site for ATP binding. Step 3: The ATP binding autophosphorylates tyrosine and activates TK.

Redrawn from (Puchner, Alexandrovich et al. 2008).

2.2 ALTERNATIVE SPLICING

Alternative splicing (AS) is a post-transcriptional process whereby a single gene encodes for multiple protein isoforms. In constitutive splicing, introns are discarded and exons are ligated in the order they are arranged in a gene. AS is a diversion from the arranged order of exons in a gene, resulting in a distinct pattern of exon-exon junctions in mature mRNA. Hence, AS is a major

source of escalating proteome diversity and its functional complexity (Liu, Gonzalez-Porta et al. 2017).

The splicing is highly influenced by the 5' splice site sequence, branch point sequence and 3' splice site sequence (Wang and Burge 2008). The branch point sequence is located few nucleotides upstream of the 3' end of intron and has a consensus sequence of "YUNAY" (Y=Cytosine/ Thymine, N=any nucleotide) where A serves as branch site and along with U is well conserved (Gao, Masuda et al. 2008). Statistical analysis of 22,489 splice junction pairs from GenBank database showed that 98.71% of splice junction sequences are 5' GT (donor site) and 3' AG (acceptor site) and are termed canonical splice sites, whereas 0.56% of splice junction pairs contain major non-canonical GC-AG as donor and acceptor sites. Lastly, the remaining 0.73% of splice site pairs contain multiple distinct non-canonical splice sites (Burset, Seledtsov et al. 2000).

Spliceosomes are large ribonucleoprotein structures that catalyze splicing. The major spliceosome is assembled from small nuclear RNAs (snRNAs) U1, U2, U4, U5, U6 and around 300 different proteins (Nilsen 2003). In addition, trans-acting splicing elements i.e. serine and arginine-rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs) have a critical role in splicing regulation (Anna and Monika 2018). One of the well-studied splicing regulators is RMB20. A study details the role of RBM20 (RNA binding motif protein 20) as a TTN splicing repressor in rats and humans. RBM20 deficiency results in unusually large titin isoform expression and is implicated in dilated cardiomyopathy (Guo, Schafer et al. 2012).

Mutations can impair canonical splice sites, generate new splice sites and activate cryptic splice sites. Moreover, the mutations in cis-regulatory elements such as exon and intron splicing enhancers and silencers may result in deviation from the normal splicing pathway (Anna and Monika 2018). To cope with aberrant splicing pathway to a certain extent, the cellular machinery has an excellent surveillance mechanism called non-sense mediated decay (NMD), which targets the mRNA transcripts for degradation if they harbor a premature termination codon (PTC). However, NMD is inefficient if the PTC is located in the last exon and the efficiency of NMD also decreases if PTC is present few nucleotides upstream of the last exon (Lindeboom, Supek et al. 2016).

It has been estimated by a compendium of splice/exon-exon junctions analyzed from RNA-Seq data that approximately 90% of all multi-exon human genes undergo alternative splicing, many in tissue and/or developmental stage-specific manner (Pan, Shai et al. 2008, Wang, Sandberg et al. 2008).

Exon skipping is the most prevalent mechanism of alternative splicing, which either eliminates or retains a particular exon from pre-mRNA. Other modes of AS are use of alternative 5´ or 3´ splice site (ASS) (partial exclusion of exon), mutually exclusive exons (one of the two adjacent exons is retained), intron retention (intronic sequence remains intact), use of alternative polyadenylation sites (different exon at 3´ of transcript) (Figure 5a) alternative promoter selection (isoforms with differing translation start sites) (Keren, Lev-Maor et al. 2010) (Figure 5b). The alternative polyadenylation mode also influences translation and isoform turnover rate based on the length of 3´UTR and poly (A) tail (Tanguay and Gallie 1996). In addition, complex alternative splicing patterns also occur in a transcript (multiple alternative splicing events [ASEs] in a single pre-mRNA molecule) that can generate thousands of splicing variants (Park, Pan et al. 2018) (Figure 5c).

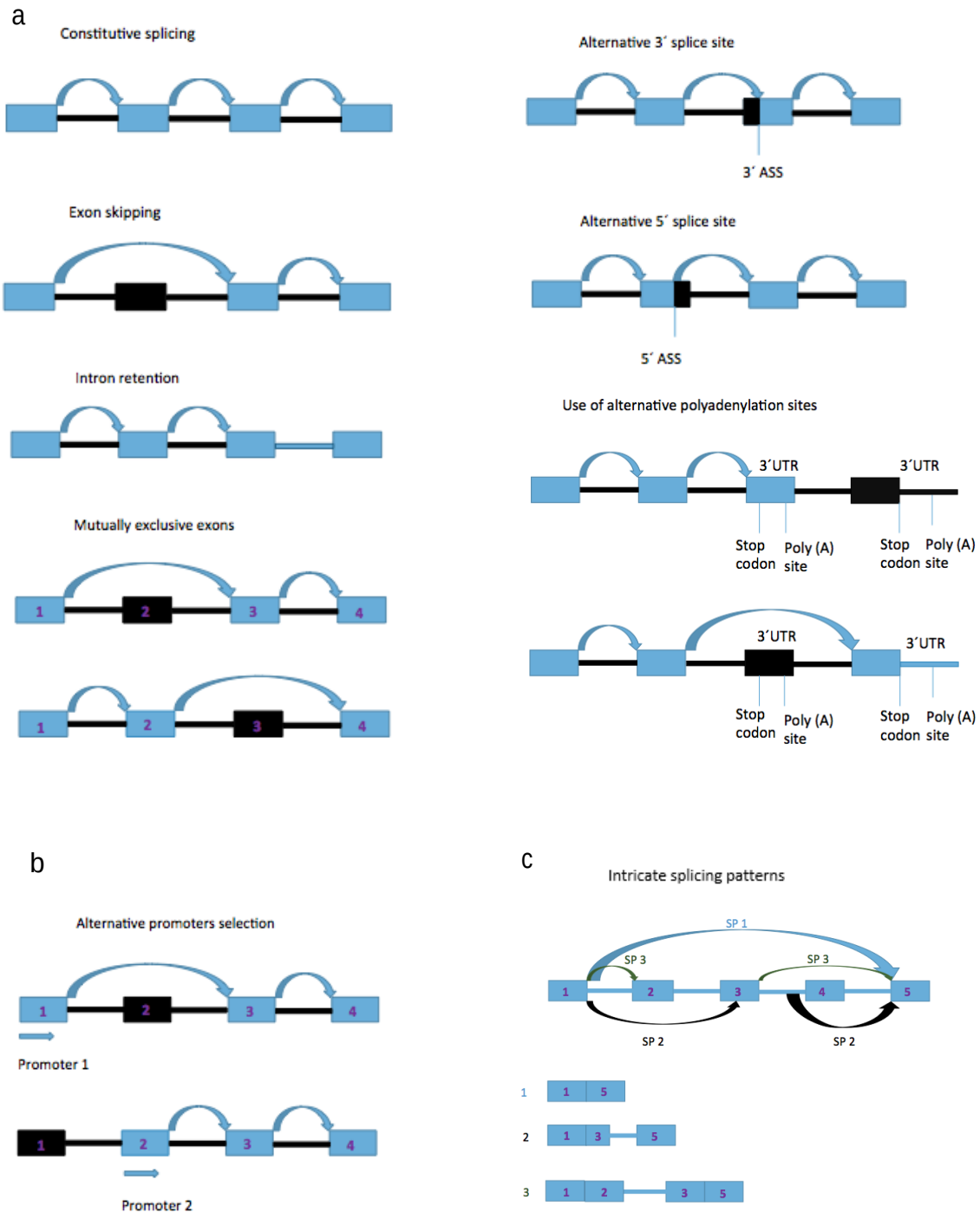

Figure 5: (a) General examples of various modes of alternative splicing. (a) and (b) Rectangles and bars colored light blue represent segments spliced in, whereas rectangles and bars colored black represent segments discarded from final transcript. (b) Alternative promoter selection is an indirect pattern of splicing; if promoter 1 is used as transcription start site, the splicing machinery has to scission the following promoter because acting as a 5' terminal exon (1st exon) of Promoter 1 it lacks 3' splice site and cannot be recognized by splicing factors. (c) Intricate patterns of splicing e.g. splicing pattern 2 (sp2) is a combination of partial exon exclusion and partial intron retention.

Redrawn from (Srebrow and Kornblihtt 2006) and (Park, Pan et al. 2018).

Since most multi-exon genes produce multiple isoforms, one of the many isoforms is generally regarded as canonical isoform on the basis of expression, length, etc. Despite being transcribed from a single gene, many isoforms are functionally disparate (Li, Menon et al. 2014). Moreover, an isoform may execute an antagonistic effect or a dominant-negative effect, i.e. two isoforms of BCL2L1 gene have opposite functions in apoptosis; the longer isoform impedes the process while the shorter isoform promotes apoptosis (Boise, Gonzalez-Garcia et al. 1993).

2.2.1 TITIN ISOFORMS AND SPLICING VARIANTS

TTN contains 364 exons, of them 363 are coding except the very first non-coding exon. This enormous number of exons in contrast to around 8-10 exons, which is an average number of exons in human genes (Sakharkar, Chow et al. 2004) enables TTN to generate thousands of splicing variants through differential splicing.

NCBI RefSeq transcript database lists seven main titin isoforms and none of them contain all 364 exons (Table 1). Titin isoforms are a result of alternative splicing principally in I-band region. In contrast, the Z-disc, A-band and-M-band sections are largely conserved among different isoforms (Labeit and Kolmerer 1995).

Generally, titin isoforms are classified on the basis of presence of skeletal muscle-specific N2A element and cardiac muscle-specific N2B element in the I-band region, and N2BA that co-expresses both elements in cardiac muscles. Furthermore, some TTN exons are skipped in all isoforms except the all-inclusive meta transcript isoform (skips exon 48 only), which is speculated to be expressed only during prenatal life (Savarese, Jonson et al. 2018).

The N2A and N2BA isoforms retain most of the I-band exons, whereas the splicing pathway in N2B isoform splices exon 50 to exon 220; thereby skipping many Ig-like domains, PEVK domains and unique sequence encoding exons of I-band (Linke 2008). The N2B isoform has less extensible potential due to its relatively shorter I-band (Cazorla, Freiburg et al. 2000) and the average expression ratio of N2BA:N2B is around 30:70. However, due to the up-regulation of N2BA in dilated cardiomyopathy the average N2BA:N2B ratio is approximately 42:58 in DCM patients which consequently lessens the passive myocardial stiffness (Makarenko, Opitz et al. 2004).

Conversely, cardiac passive stiffness increases during postnatal development when fetal cardiac titin expression is superseded by adult cardiac titin isoforms (Lahmers, Wu et al. 2004).

Three additional cardiac titin isoforms are produced via mutually exclusive splicing mode, namely Novex 1 (exon 45), Novex 2 (exon 46) and Novex 3 (exon 48) (Bang, Centner et al. 2001). The Novex 1 and Novex 2 are identical to N2B isoform with the additional inclusion of 125 (Exon 45) and 192 (Exon 46) unique amino acid sequences in I-band region, respectively (Savarese, Sarparanta et al. 2016). The 'tiny titin' Novex 3 is the only isoform that retains exon 48, which contains alternate termination codon resulting in proportionately small 700 kDa protein with alternative C-terminal (Bang, Centner et al. 2001). Moreover, an additional isoform called Cronos that is produced with an alternative promoter from I-band has also been reported and linked as a likely contributor to severe cardiomyopathy phenotype (Deo 2016).

RNA-Seq data obtained from left ventricle biopsies of DCM patients and anatomically different skeletal muscle biopsies showed that some exons are constitutively expressed while few exons are inconsistently incorporated in a particular isoform (Roberts, Ware et al. 2015, Savarese, Jonson et al. 2018). This differential exon usage is measured as PSI value (percent/proportion spliced in) that estimates the mean value of a specific exon inclusion in mature transcripts; the higher PSI refers to frequent exon usage and vice versa (Roberts, Ware et al. 2015).

To some extent, TTN structure is appropriate for extensive alternative splicing as 85% of the TTN exons are symmetric (entire exon sequence is in triplets) and their removal does not render frameshift effect. To a certain extent the PSI corresponds to exon symmetry, for instance, 49 of 185 exons with PSI >99% are asymmetric. Furthermore, the lowest PSI value is observed in TTN I- band region where 93% of alternatively spliced exons are symmetric (Roberts, Ware et al. 2015). The higher PSI in asymmetric exons is likely to be favorable owing to the fact that asymmetric exon exclusion can distort translation reading frame.

Table 1: The table of seven main human TTN isoforms.
(Modified from https://www.cardiodb.org/titin/titin_transcripts.php and updated according to NCBI nucleotide database)

Isoform	Description	Transcript Length (Excluding 5´ cap and poly (A) tail)	Polypeptide Length	Number of Exons	Refseq Transcript	Refseq Protein
Meta	Complete meta transcript	109224	35991	363	NM_001267550.2	NP_001254479.2
N2BA	Principle cardiac long isoform	104301	34350	313	NM_001256850.1	NP_001243779.1
N2A	Soleus/ skeletal long isoform	101520	33423	312	NM_133378.4	NP_596869.4
N2B	Principle cardiac short isoform	82029	26926	191	NM_003319.4	NP_003310.4
Novex 1	Minor cardiac short isoform	82404	27051	192	NM_133432.3	NP_597676.3
Novex 2	Minor cardiac short isoform	82605	27118	192	NM_133437.4	NP_597681.3
Novex 3	Minor cardiac short isoform	18236	5604	46	NM_133379.4	NP_596870.2

2.3 TITINOPATHIES

Titinopathy/Titin myopathy is an umbrella term that encompasses wide-ranging skeletal muscle disorders (Figure 6) and cardiac muscle disorders linked to mutations in TTN gene. Titinopathies are inherited in both autosomal dominant and recessive fashions. Some well-characterized titinopathies are late-onset dominant tibial muscular dystrophy, young/early adult-onset recessive distal titinopathy, limb-girdle muscular dystrophy type 2J, congenital centronuclear myopathy and early-onset myopathy with fatal cardiomyopathy (Savarese, Sarparanta et al. 2016). Furthermore, dominant De Novo mutation in TTN, leading to cardiac failure has also been documented (Peled, Gramlich et al. 2014).

The first TTN mutation (arg to leu replacement in codon 740) was reported in 1999; the authors proposed that the mutation has a high likelihood of being the underlying cause of hypertrophic cardiomyopathy (Satoh, Takahashi et al. 1999). Though, the first confirmed association of TTN mutation to a skeletal muscle disorder was published in 2002; the mutation was described as 11 bp change (AAGTAACATGG→TGAAAGAAAAA/EVTW→VKEK, FINmaj mutation) in the last exon of TTN, observed in 81 Finnish patients (Hackman, Vihola et al. 2002).

2.3.1 MISSENSE TITINOPATHIES

One of the well-known missense titinopathies is hereditary myopathy with early respiratory failure (HMERF) that is characterized by debilitating muscles of lower and upper extremities and affecting respiratory muscles too. The pathogenic mutations in HMERF are located in a single specific exon (344) and act in a semi-dominant way (homozygous patients have more severe condition than heterozygous patients) (Palmio, Leonard-Louis et al. 2019).

2.3.2 NONSENSE/TRUNCATED/SPLICING TITINOPATHIES

The term titin truncating variant (TTNtv) can be misapprehended. The literal definition of truncating variant denotes any change that shortens the normal size of the encoded protein; however, TTNtv is commonly used to describe TTN splicing, nonsense and frameshift variants, even though their post-transcriptional outcomes may differ greatly.

In context of skeletal muscle disorders, all TTNtvs are recessive (Ceyhan-Birsoy, Agrawal et al. 2013, Chauveau, Bonnemann et al. 2014, Peric, Glumac et al. 2017), however, in many DCM cases

TTNtvs seem to act in a dominant fashion (Tabish, Azzimato et al. 2017) . The TTNtvs are prevalent in DCM cases; approximately 20% of DCM patients have TTNtvs enriched in the A-band region where all exons are consistently included in the mature transcript (Herman, Lam et al. 2012).

Interestingly, around 2% of the healthy human population has TTNtvs that are likely to be phenotypically silent possibly because TTNtvs observed in control samples were enriched in exons not utilized in cardiac N2BA and N2B isoforms (Roberts, Ware et al. 2015), suggesting the penetrance of causative TTNtvs in general population is significantly low.

It is a well-established genetic fact that the relative abundance of transcripts can be a key factor in disease manifestation in comparison to a healthy state; such quantitative evaluation has not been formally done for TTNtvs. Nevertheless, an estimated 0.5% of the general population carries TTNtvs in exons with PSI >15% that may have the potential to cause DCM (Schafer, de Marvao et al. 2017).

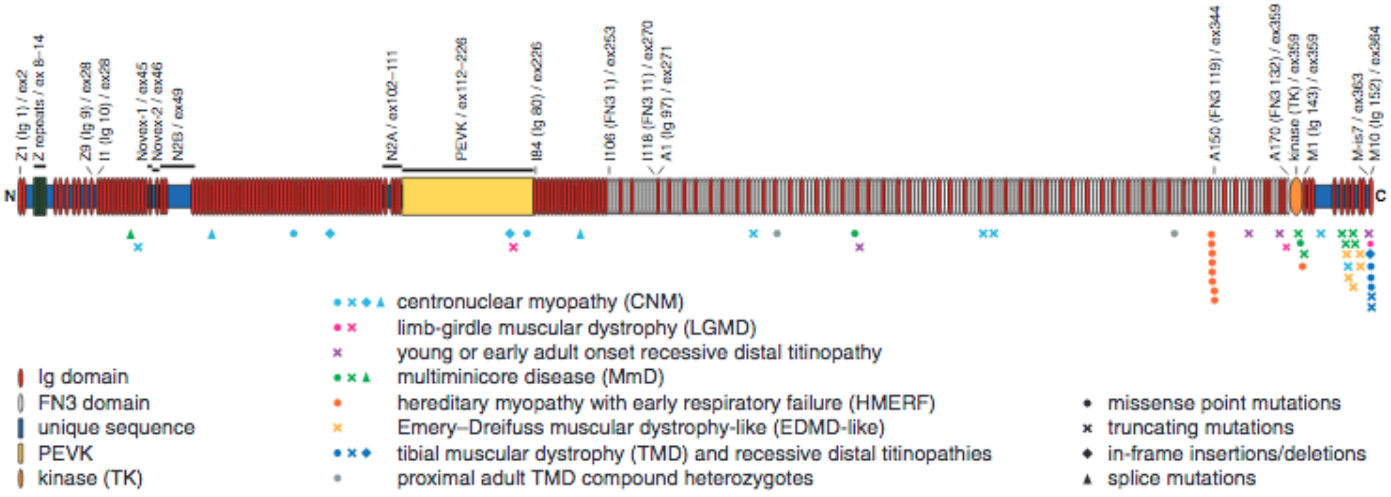

Figure 6: The figure exemplifies well-known skeletal muscle disorders linked to TTN mutations. The symbols shapes and colors denote mutation type and the associated phenotype, respectively. The symbols also mark the relative positions of mutations in particular domains.

Adapted from (Savarese, Sarparanta et al. 2016) with permission from authors.

3 AIMS OF THE STUDY

This thesis serves two purposes.

1) Characterization of TTN splicing variants to enhance the diagnostic accuracy for patients with skeletal muscle disorders and to further elaborate on pathogenesis of titinopathies.
2) An attempt to resolve which sequencing method is most suitable for TTN splicing variants characterization.

4 MATERIALS AND METHODS

4.1 PATIENTS' SAMPLES

This study included six patients of confirmed titinopathy i.e. biallelic TTN truncating variants (detected by NGS MyoCap assay) and a phenotype compatible with the previously reported titinopathies.

Individuals without any skeletal muscle disorder were recruited as a control group.

All individuals (control subjects and patients) enrolled in this project were thoroughly informed about the study and have provided written consent. The ethics of the research have been reviewed by the ethics committee of the Helsingin ja Uudenmaan sairaanhoitopiiri (HUS, statement number 195/13/03/00/11).

The study has been performed according to the guidelines of The Declaration of Helsinki.

4.2 METHODS

4.2.1 RNA EXTRACTION AND SEQUENCING

Total RNA was extracted and purified from skeletal muscle biopsies using QIAGEN RNeasy Plus Universal Mini Kit. Deep frozen muscle samples (10-20mg) were placed in lysis tubes E (Analytik Jena AG) containing steel beads and immediately added 900µl QIAZol lysis reagent. The samples were homogenized on Tissuelyser (SpeedMill PLUS Analytik jena AG) for 2-5 minutes at 20 Hertz. The subsequent steps were conducted according to the protocol (QIAGEN RNeasy Plus Universal Handbook Dec 2014). The optional step of column drying was performed and RNA was eluted in 40µl RNAse free water with second elution using the eluate.

RNA was quantified using Denovix DS-11 Spectrophotometer and qualitative assessment was done using High Sensitivity RNA ScreenTape (Agilent Technologies) on Agilent 4200 TapeStation system (Agilent Technologies). The samples were dispatched to The Oxford Genomics Centre for full

transcriptome analysis (after ribodepletion). RNA enrichment was done using Illumina's Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat) and with Illumina´s TruSeq Stranded Total RNA Library Prep (Human/Mouse/Rat), following the manufacturer´s instructions.

Library preparations underwent a full quality control evaluation and were sequenced with HiSeq4000 (read length=75bp, paired-end run). An average of ~80 million reads was produced across all samples. Adapter sequences and low quality bases were removed with fastp. Trimmed sequences were then mapped with STAR 2.7.0d with index generated from Gencode.v29 human reference (release date 05.2018, based on ENSEMBL GRCh38.p12).

4.2.2 REVERSE TRANSCRIPTION

The extracted RNA was reverse transcribed as written below in tables 2 and 3.

Table 2: The master mix composition of reverse transcription reaction

Reagent	Quantity
10xRT Buffer	2µl
10x Random Primers	2µl
25x DNTP Mix (100 mM)	0.8µl
~100ng/µl RNA Template	10µl
Nuclease-free Water	4.2µl
MultiScribe Reverse Transcriptase	1µl
	Total 20µl

Table 3: The thermal conditions and their relative timings for reverse transcription

Phase	Temperature	Time
Annealing	25°C	10 Minutes
Extension (cDNA synthesis)	37°C	120 Minutes
Reverse-Transcriptase deactivation	85°C	5 Minutes
Incubation	4°C	∞

4.2.3 POLYMERASE CHAIN REACTION

The PCR reactions were performed on reverse transcribed cDNA. Primers were designed on Primer3 software (version 4.1.0, http://primer3.ut.ee/) and ordered from SIGMA-ALDRICH. When possible (if exon-exon junctions were known), exon-exon junction primers were used for PCR amplification to circumvent genomic DNA synthesis. The quantities of reagents and timings of steps are as follows (Tables 4 and 5). The DNA concentrations were checked spectrophotometrically on Nanodrop (DeNovix DS-11 FX +).

Table 4: The master mix composition of cDNA synthesis

Reagent	Quantity
5x GC Buffer	6µl
10mM DNTPs	0.6µl
cDNA	0.5µl
10 µM Forward Primer	1µl
10 µM Reverse Primer	1µl
Milli-Q Water	20.5µl
2U/µl Phusion Hot Start II DNA Polymerase	0.4µl
	Total 30µl

Table 5: The thermal conditions and their relative timings for cDNA synthesis in 36 cycles

Phase	Temperature	Time
Initial Denaturation	98°C	30 Seconds
Denaturation	98°C	10 Seconds
Annealing	60°C	30 Seconds
Elongation	72°C	3 Minutes
Additional 35 cycles of Denaturation, Annealing and Elongation.		Approximately 130 Minutes.
Final Elongation	72°C	10 Minutes
Incubation	4°C	∞

4.2.4 GEL DNA EXTRACTION

(1.5% agarose gel prepared in 0.5x TBE with 5µl of midori green advance [NIPPON Genetics EUROPE] added)

20µl PCR amplified products intermixed with 4µl of 6x loading dye (Thermo Fischer Scientific) were run on 1.5% agarose gel submerged in 0.5X TBE Buffer at 110V for 45-60 minutes. The bands visualized under UV light were excised and DNA was extracted following the manufacturer's protocol (MACHEREY-NAGEL NucleoSpin Gel and PCR Clean-up February 2017/Rev. 04) with the recommended step of heating elution buffer at 70°C. The purified DNA concentration was checked on Nanodrop (DeNovix DS-11 FX +). The purified DNA was Sanger sequenced by the same primers used for amplification. The sequences were checked on Sequencher software (Gene Codes Sequencher 5.3).

4.2.5 PCR CLONING

The PCR fragments were ligated into PCR blunt II-Topo vector (Invitrogen Zero Blunt TOPO PCR Cloning Kit). The ligation mixture was prepared using 1.5µl PCR product, 0.5µl salt solution, 0.5µl Blunt II-Topo vector and 0.5µl water with subsequent room temperature incubation for 5 minutes. Ice thawed vial of chemically competent cells (One Shot TOP10 Chemically Competent E. coli) was pipetted into the ligation reaction followed by 30 minutes incubation on ice (to attach cloned vector to the bacterial membrane).

The sample was then heat shocked for 30 seconds at 42°C (increases membrane permeability to facilitate the transport of vector into cells) and instantly set on ice for 2 minutes (closes membrane to avoid leaking of cell contents). To obtain maximum transformation efficiency, the cells were nourished with 250µl of SOC media (Thermo Fischer Scientific) (culturing in antibiotic-free media before plating allows expression of antibiotic resistance gene) and put on thermo shaker at 300 RPM (revolutions per minute) and 37°C for one hour. The SOC grown cells were plated on ampicillin (100µg/µl) agar plates and subjected to overnight incubation at 37°C.

4.2.5.1 OVERNIGHT BACTERIAL CULTURE

The following day 5 sparsely grown colonies were picked and each colony was inoculated in 5ml L.B media containing 500µg/5µl of ampicillin and allotted another overnight incubation at 37°C while shaking at 300 RPM under aerobic condition.

4.2.5.2 PLASMID ISOLATION

The plasmids were isolated as detailed in the user manual (NucleoSpin plasmid DNA purification/Isolation of high-copy plasmid MACHEREY-NAGEL December 2015/Rev. 09). No recommended steps were performed and plasmids were eluted in 40µl water instead of buffer AE. The purified plasmid concentration was checked on Nanodrop (DeNovix DS-11 FX +) and sent for DNA sequencing as per section 4.2.6, with the exception of ExoSAP purification.

4.2.6 SANGER SEQUENCING

The Sanger sequencing was done to ascertain NGS MyoCap detected mutations and on PCR amplified cDNA. Prior to sequencing, 3µl of ExoSAP-IT was added to 10µl of PCR product followed by incubation at 37°C (to digest excessive primers and free nucleotides) for 30 minutes and then at 80°C (to inactivate ExoSAP-IT to avoid its action on primers and nucleotides during sequencing) for 15 minutes.

The sequencing reaction was prepared with 1-4µl of DNA (depending on the size of the template), 1.5µl (10 µM) of either forward or reverse primer keeping the final volume to 6µl by adding water. The samples were sequenced at Institute for Molecular Medicine Finland. The sequences were analyzed on Sequencher software (Gene Codes Sequencher 5.3).

4.2.7 SINGLE-MOLECULE REAL-TIME SEQUENCING (LONG READS)

The PCR amplified cDNA samples covering particular TTN exons regions and 30 other muscle genes cDNAs amplified from 5'UTR to 3'UTR region were normalized according to their relative concentrations and fragments size, and pooled into a single tube. A size selection was performed to remove amplicons shorter than 700bp to circumvent the frequent sequencing of shorter fragments over large ones. Library preparation was performed at DNA sequencing and genomics facility at the Institute of Biotechnology, Helsinki Institute of Life Science (HiLIFE), University of

Helsinki. The final library was run on PacBio RSII. After trimming the adapters, sequences were manually visualized and inspected using Integrative Genomics Viewer (IGV).

5 RESULTS

All TTN variants are described using Meta transcript NM_001267550.1, according to the Human genome variation society's (HGVS) recommendations. Exon numbering is according to the Leiden open variation database (LOVD), i.e. LRG numbering, which is recommended for clinical reporting and does not correspond to the RefSeq numbering.

5.1 COMPOUND HETEROZYGOSITY

NGS MyoCap assay unveiled the presence of bi-allelic TTN mutations in all patients included in this study. Most TTN previously unreported variants and VUS were single nucleotide variants (SNVs), detected in canonical splice sites. In addition, Indel and nonsense mutations were also detected with variants in canonical splice sites (compound heterozygous state) (Table 8).

All patients had at least one variant in either donor or acceptor splice site. In particular, patients 6 and 1 were identified with bi-allelic variants in canonical splice sites. Patients 2, 5 and 4 were detected with small insertion or deletion in a compound heterozygous state with a splicing variant. Lastly, patient 3 was also detected with a splicing variant and a nonsense variant in a compound heterozygous state. The additional 235 muscle-related genes sequenced in MyoCap were devoid of any potential causative mutation responsible for the observed phenotype.

The outline of clinical phenotypes of patients is given in table 6, and the methods used for splicing variants characterization and their results are recapitulated in table 8.

Table 6: Salient clinical features of patients

Patient	Sex/origin	Age at last examination/age of onset	Symptoms at onset	Muscle weakness	Phenotype
1	F/France	-/Early childhood	Difficulty in walking	Proximal> LL distal	Early onset recessive titinopathy
2	M/Finland	-/Birth	Hypotonia and arthrogryposis	Distal LL and neck flexor	Congenital myopathy
3	M/North America	14/Birth	Hypotnia	Distal and proximal	Congenital myopathy
4	M/Belgium	34/Infancy	Progressive muscular weakness and gait abnormalities accompanied by frequent falling.	Proximal UL, proximal and distal LL	Congenital myopathy
5	M/Spain	47/Early adulthood)	Distal LL weakness and mioglobinuria	Psoas and gluteii 4/5; partial atrophy of quadriceps, addcutors 3/5, tibialis anteior 4/5, hanging big toe	Distal myopathy
6	M/Italy	57/Adulthood	Difficulty in walking	Distal>proximal	Distal myopathy

5.2 RESULTS FROM TRADITIONAL SANGER SEQUENCING METHODS

Patient 1

MyoCap evidenced two previously undescribed variants in this patient: c.15776-1G>T (intron54) and c.67349-2A>C (intron 319).

PCR reactions were performed using two different combinations of primers (exons 54-56, exons 54-57) on patient and control cDNA samples. The gel electrophoresis analysis confirmed the amplification of fragments of expected size (approximately 200bp and 250bp) that were also seen from control samples. The additional larger band visible only from the patient's sample suggested a splicing effect due to the abovementioned heterozygous variant.

Direct Sanger sequencing results were enriched of ambiguous base calls, hampering correct analysis of splicing; thus, the PCR product was rerun on agarose gel electrophoresis, and the outlined DNA bands were gel-extracted (Figure 7). The subsequent Sanger sequencing revealed 96 nucleotides long intron 54 retention (Figure 8). A similar result was also observed from the sequencing of cloned PCR product.

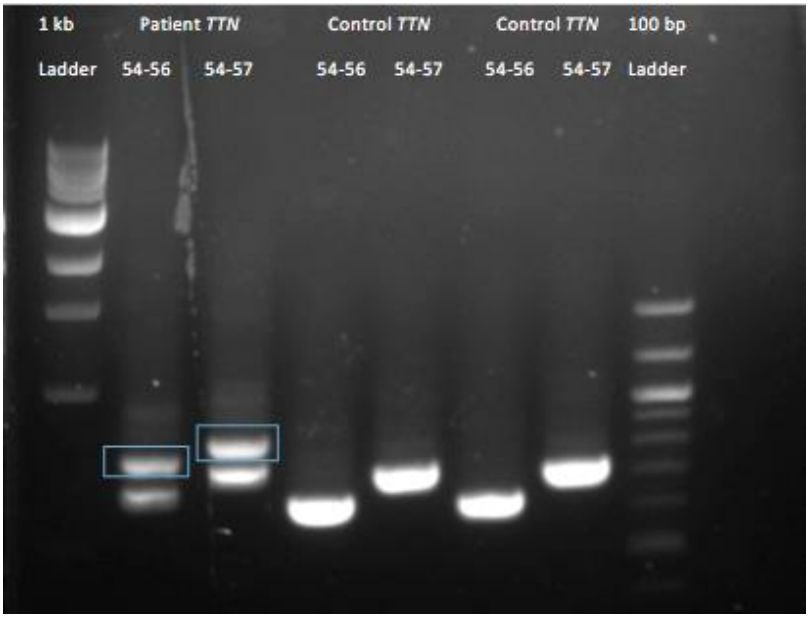

Figure 7: The PCR amplification of exon 54 to either 56 or 57. Two bands appear on patient samples show the heterozygous state, the presence of larger fragments with reference to control fragments is an indication of intronic retention/splicing defect.

Direct Sanger sequencing and Gel-extracted DNA sequencing were also performed to characterize the variant c.67349-2A>C. However, the results were elusive due to low signal to noise ratio. The cloning of the PCR products (amplified using primers on exons 319-321) with subsequent Sanger sequencing led to the identification of two discrete splicing effects, entire intron 319 retention (Figure 9) and first 15 nucleotides loss of exon 320 (Figure 10)

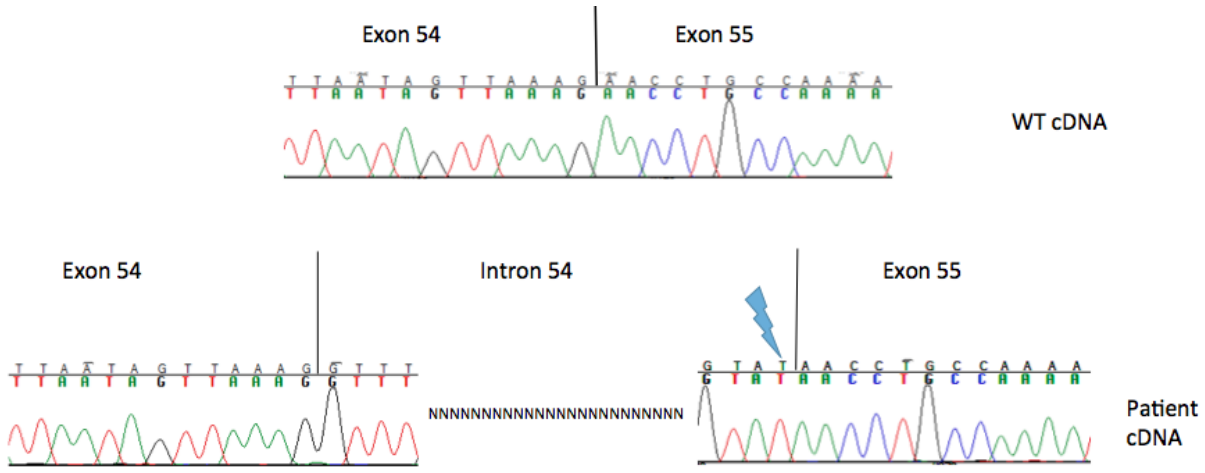

Figure 8: The chromatogram shows mutation c.15776-1G>T in intron 54. The downstream effect is entire 96 nucleotides long intron 54 retention.

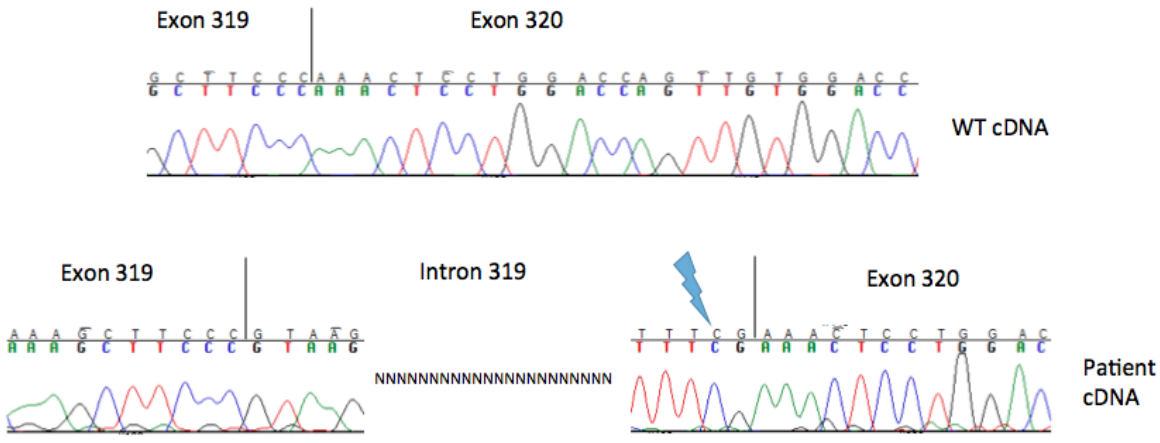

Figure 9: The chromatogram shows mutation c.67349-2A>C in intron 319. The downstream effect is entire intron 319 retention.

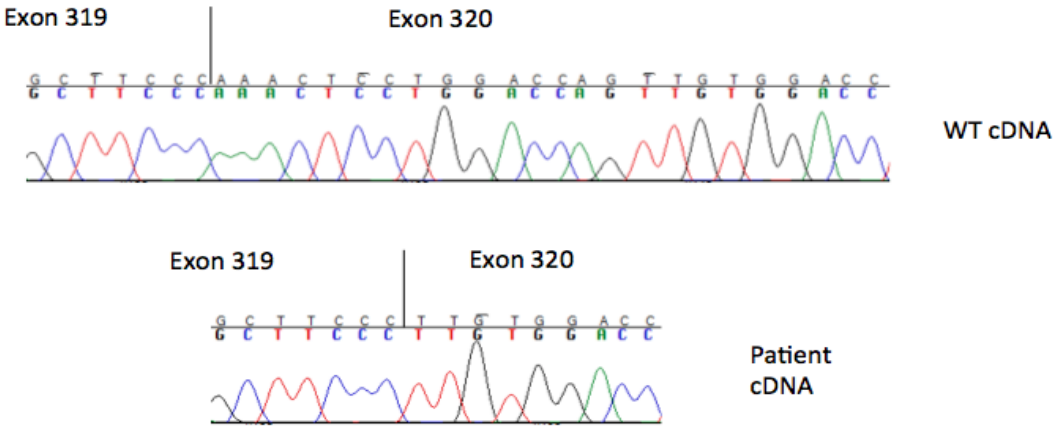

Figure 10: The aforementioned mutation executes a different splicing effect resulting in loss of upstream 15 nucleotides of exon 320 and p.(Gln22450Leu).

Patient 2

Patient 2 carries a variant affecting the canonical splice site (c.37040-1G>A intron 179) and 4 nucleotides deletion in exon 358 c.100312_100315del p.(Lys33438Glyfs*4).

Only the variant in intron 179 was selected for characterization because the other variant in exon 358 was not anticipated to alter splicing.

Patient cDNA was PCR amplified with two pairs of primers on different exons (168-182 and 171-182). The results drawn after PCR cloning and sequencing were inconclusive because exon 179 is meta-only exon (not expressed in adult muscles) and located in the repetitive region. The reasons are further elaborated in discussion.

Patient 3

Patient 3 was identified with SNV in allele 1 c.53354G>A exon 279 p.(Trp17785*) that renders a nonsense effect. The consensus splice site variant in other allele was c.19426+2T>A localized in intron 67.

PCR amplified fragments spanning the region from exon 64 to 68 were cloned and sequenced. Two distinct exon skipping events were seen on chromatogram, the one excluding exon 67 and other less frequent excluding exon 66 and 67.

Patient 4

From MyoCap, two variants were identified c.35829_35830insA p.(Gly11944Argfs*7) in exon 164 and c.25063+1G>A in intron 87. The splicing variant was selected for sequencing.

PCR amplified fragments spanning the region from exon 85 to 90 were Sanger sequenced, and the skipping of exon 87 was observed.

Patient 5

Along with a single nucleotide deletion in exon 364 c.107889delA p.(Lys35963Asnfs*9), A variant affecting the canonical splicing site was observed in intron 26 (c.4646-1G>A). A PCR spanning the region from exon 25 to 28 was performed. A direct Sanger sequencing showed the skipping of first 14 nucleotides of exon 27.

Patient 6

MyoCap identified two variants in canonical splice sites, in particular, c.107377+2dup in intron 362 and c.54190+1G>A in intron 281.

A PCR (primers in exon 361 and 364) followed by Sanger sequencing revealed the retention of the first nucleotide of intron 362.

Similarly, a PCR (primers in exons 280 and 283) was performed on cDNA of patient 6 and control. The most intense band only seen in the patient was extracted, and the subsequent Sanger sequencing showed the retention of 88 nucleotides long intron 281.

5.3 RESULTS FROM SECOND-GENERATION (ILLUMINA RNA-Seq) AND THIRD-GENERATION (SMRT) SEQUENCING APPROACHES

Two innovative techniques were implemented to examine the splicing effects. Every single splicing variant was studied using total RNA-Seq after Ribo-depletion based library preparation, whereas all except one splicing variant (c.54190+1G>A) were studied using SMRT due to the lack of sample.

For total RNA-Seq, we analyzed the junction files and manually inspected the TTN region of interest using IGV.

RNA-Seq confirmed all the previously detected splicing effects from traditional Sanger sequencing. As expected, the intronic retentions in patients 1 and 6 were clearly seen in IGV due to the presence of reads in introns 319 and 281, respectively. Conversely, no intronic reads were seen in samples without a possible splicing variant in these TTN regions. Similarly, no reads provided coverage on other TTN introns in patients 1 and 6, proving that reads in introns 319 and 281 are due to their retention and not because of possible DNA contamination.

Interestingly, in patient 4, additional splice events, not identified by the traditional PCR and Sanger sequencing, were identified. The retention of intron 87 was clearly observed in IGV (Figure 11). No reads in other TTN introns were observed in this patient, and no other samples showed a coverage in intron 87, demonstrating the clear mutational effect. Moreover, the skipping of exon 87 alone and with exon 86 (junction 85-88,) was observed with 1353 and 1536 reads, respectively.

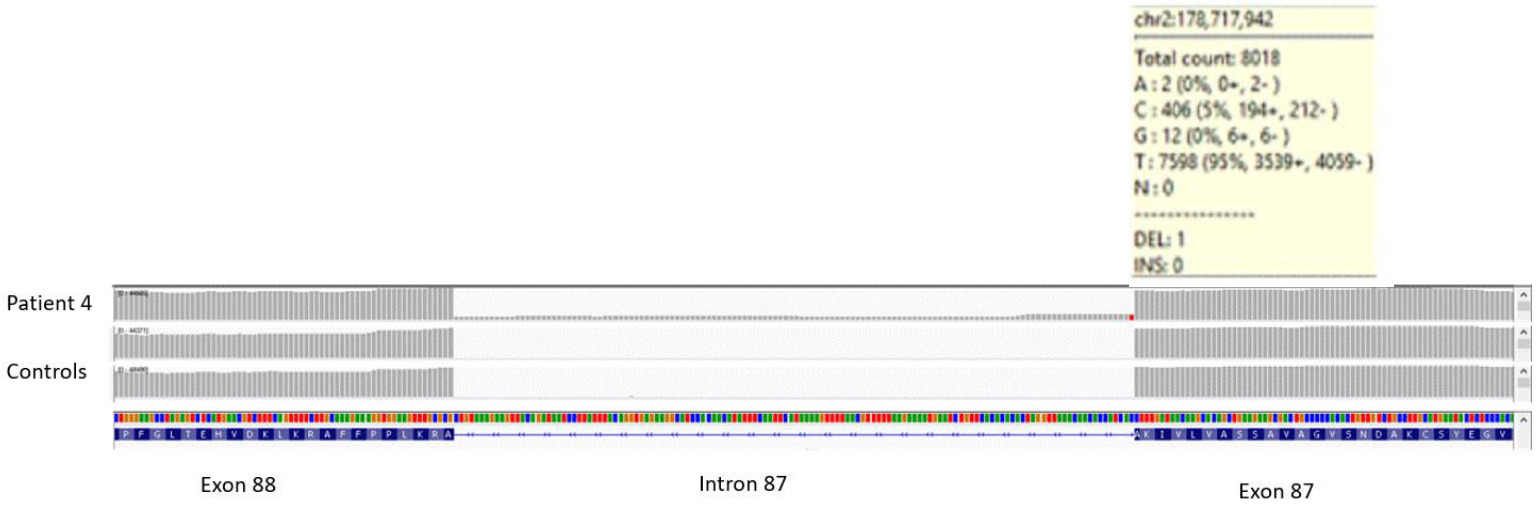

Figure 11: The figure from integrative genomics viewer showing intron 87 retention in patient 4. Out of total 8018 reads, 7598 reads support the variant call c.25063+1G>A (In figure shown by nucleotide T because of TTN location on minus strand).

Similarly, in patient 6 two additional abnormal splicing events due to the variant c.54190+1G>A in intron 281 were observed; the loss of last 7 and last 130 nucleotides of exon 281 was supported by 966 and 1890 reads, respectively.

By further analysis of exon junction files, we also tried to quantify aberrant splicing events of exon skipping and loss of nucleotides (Table 7).

In patient 1, 23317 reads supported the loss of upstream 15 nucleotides of exon 320 and were more than twice the number of reads supporting exon 319-320 junction, which supposedly came from wild type allele.

In patient 3, the skipping of exon 67 (a junction 66-68) was explicitly observed with 14648 reads from RNA-Seq results. Reads supporting exclusion of exon 66 along with exon 67 (junction 65-68) were only 544 in number. In contrast, 22001 reads supported the exon 66-67 junction.

In patient 5, the loss of first 14 nucleotides of exon 27 was confirmed by 1183 reads, the number is approximately 1/3 of reads supported exon 26-27 junction (Table 7).

Finally, in patient 6, 10334 reads from RNA-Seq results supported the retention of the first nucleotide of intron 362 from IGV (10312 supported the normal junction 362-363).

SMRT analysis confirmed the previous PCR-based results but, for example, failed to identify the skipping of exon 87 and 86 along with 87 in patient 4 where only the intronic retention was detected. The table below enumerates RNA-Seq reads from splicing variants.

Table 7: RNA-Seq Reads (Exon-exon junctions and loss of nucleotides, reads of interest are written in bold)

Patient	Splice variant	Reads	Donor exon	Acceptor exon	Genomic coordinates
1	c.67349-2A>C (int319)	23317	319	#N/A	2:178579833-178579939 (loss of first 15 nucleotides of exon 320)
		9570	319	320	2:178579848-178579939
2	c.37040-1G>A (int179-metaonly) undetected	12	179	189	2:178657767-178662142
		19	179	180	2:178662028-178662142
3	c.19426+2T>A (int67)	16145	67	68	2:178728397-178728500
		14648	66	68	2:178728397-178728891
		22001	66	67	2:178728778-178728891
		544	65	68	2:178728397-178729288
4	c.25063+1G>A (int87)	30470	87	88	2:178717810-178717943
		42988	86	87	2:178718221-178718322
		1353	86	88	2:178717810-178718322
		1536	85	88	2:178717810-178718695
		43019	85	86	2:178718600-178718695
5	c.4646-1G>A (int26)	3355	26	27	2:178777317-178777420
		1183	26	#N/A	2:178777303-178777420 (Loss of the first 14 nucleotides of exon 27)
6	c.54190+1G>A (int281)	23781	281	282	2:178604898-178604987
		966	#N/A	282	2:178604898-178604994 (loss of last 7 nucleotides of exon 281)
		1890	#N/A	282	2:178604898-178605117 (loss of last 130 nucleotides of exon 281)

Table 8: A concise description of splicing variants and their results obtained from different methods.

Patient	Splicing Variant(s)	Results	Methods	Additional Information
1	c.67349-2A>C (Intron 319)	Intron 319 retention, loss of first 15 nucleotides of exon 320.	Direct Sanger sequencing, Sanger sequencing on Gel extracted DNA bands, Sanger sequencing on cloned PCR products, RNA-Seq and SMRT sequencing.	
	c.15776-1G>T (Intron 54)	Intron 54 retention.		
2	c.37040-1G>A (Intron 179)	Inconclusive (Meta only).	Direct Sanger sequencing, Sanger sequencing on cloned PCR products, RNA-Seq and SMRT sequencing.	RNA-Seq data showed negligible number of reads supporting exon 179 inclusion in adult isoform (N2A). In trans variant c.100312_100315del
3	c.19426+2T>A (Intron 67)	Skipping exon 67, skipping exon 66 with 67.	Sequencing on cloned PCR products, RNA-Seq and SMRT sequencing.	In trans variant c.53354G>A
4	c.25063+1G>A (Intron 87)	Skipping exon 87, skipping exon 86 with 87 and intron 87 retention.	Direct Sanger sequencing, RNA-Seq and SMRT sequencing.	Skipping of exon 86 with 87 and intron 87 retention were not seen in Sanger sequencing and no skipping event was identified from SMRT. All splicing events were detected from RNA-Seq. In trans variant c.35829_35830insA
5	c.4646-1G>A (Intron 26)	Loss of the first 14 nucleotides of ex 27.	Direct Sanger sequencing, RNA-Seq and SMRT sequencing.	In trans variant c.107889delA
6	c.54190+1G>A (Intron 281)	Intron 281 retention, loss of last 7 and last 130 nucleotides of exon 281.	Direct Sanger sequencing, Sanger sequencing on Gel extracted DNA bands, RNA-Seq and SMRT sequencing.	Loss of last 7 and last 130 nucleotides of exon 281 was only observed from RNA-Seq results.
	c.107377+2dup (Intron 362)	Retention of first nucleotide of intron 362.		

6 DISCUSSION

6.1 IMPORTANCE OF CHARACTERIZING SPLICING VARIANTS

Alternative splicing is a natural phenomenon to produce multiple gene isoforms/splice variants by a single gene. The transcripts if diverge from normal splicing patterns may prompt an abnormal function or undergo degradation via NMD in certain circumstances (Ge and Porse 2014).

Although, the enhanced efficiency of genetic tests as a result of the introduction of massively parallel sequencing technologies in clinical diagnosis is phenomenal; the number of variants of uncertain significance has also increased substantially (Di Resta, Galbiati et al. 2018).

The characterization of titin splicing variants poses a challenge primarily due to its sheer size, repetitive regions, meta-only exons and highly modular structure (Savarese, Maggi et al. 2018). However, the characterization of variants is of profound importance in order to make an accurate diagnosis, especially in patients manifesting less distinct phenotypes with a possibility of numerous candidate muscle genes involved.

Furthermore, the characterization of variants is essential to distinguish between benign and causative variants to their relative extents (semi-dominant), which eventually could be useful for personalized medicine approach as well and to limit the number of variants of uncertain significance.

In this project, we studied six patients of skeletal muscle disorders, identified with TTN potential causative mutations in NGS MyoCap assay. The downstream effects of splicing were analyzed with different sequencing techniques, including third-generation sequencing called SMRT long reads, which provides an average read length of 10-16 kb (Ardui, Ameur et al. 2018).

6.2 THE COMPLEXITY OF ALTERNATIVE SPLICING PATTERNS IN TTN GENE

Due to the vast number of exons in TTN gene, exon-exon junctions that are not part of adult canonical isoforms are commonly observed and may not necessarily indicate aberrant splicing, especially in case of meta-only exons that are frequently skipped as part of normal splicing

pathway in adult isoforms. This draws attention towards the fact that whether a particular exon skipped in a patient is commonly skipped in healthy population or is a consequence of splicing defect. The number of reads supporting specific exon exclusion/inclusion in healthy population in comparison to a patient as a result of splicing variant may explain the splicing defect, (RNA-Seq data provides a good semi-quantitative estimation of the different splicing events). Besides that, junctions entering and exiting a particular exon are also often different (exon usage).

6.3 INTERPRETATION OF RESULTS

TTN splicing mutations detected in a compound heterozygous state (second allele with a different causative TTN mutation) were analyzed.

In patient 1 first splicing variant c.15776-1G>T (Intron 54) causes retention of intron 54, which has multiple stop codons and as a consequence mRNA is likely to be degraded via NMD pathway. This explanation is corroborated by previous data (Schafer, de Marvao et al. 2017), which describes that most TTNtvs are degraded by NMD. Thus, this will effectively be a null allele. Nonetheless, a small amount of truncated protein synthesis is expected given the fact that the NMD pathway is elicited following the first round of translation (Plant, Wang et al. 2004).

The second variant c.67349-2A>C (Intron 319) prompts two distinct ASEs (skipping of upstream 15 nucleotides of exon 320 and retention of intron 319). Since intron retention also triggers NMD because of PTCs occurrence within its 90 nucleotides long sequence; the skipping event (loss of one FNIII domain) would result in a near full-length titin and is apparently the reason of affected individual's survival as well the reason of disease phenotype.

For the second variant, the RNA-Seq results present more reads of 15 nucleotides skipping effect, apparently splicing machinery on this variant frequently uses cryptic splice site in exon 320 before splicing exon 320 to 321. (2262 reads showed the intronic retention from IGV [data not shown here] versus 23317 reads supported the loss of 15 nucleotides).

In the patient 2, the variant c.37040-1 G>A (Intron 179) was detected in NGS MyoCap assay but the RNA and cDNA sequencing results were not decisive because the exon 179 is only expressed in

meta-transcript (during embryonic life) and has extremely low or no expression in any adult main isoform.

PCR amplification of cDNA (if there is any expression of exon 179) using normal primers is difficult because identical sequences are present in exons 188 and 197, also junction primers are difficult to design in this particular region because junctions of flanking exons are not known as they are also meta-only exons and share repeated sequences with exons in vicinity (Figure 12). Moreover, if we assume that variant causes intron 179 retention, PCR amplification in this region would still be problematic because introns 179,188 and 197 also have identical sequences.

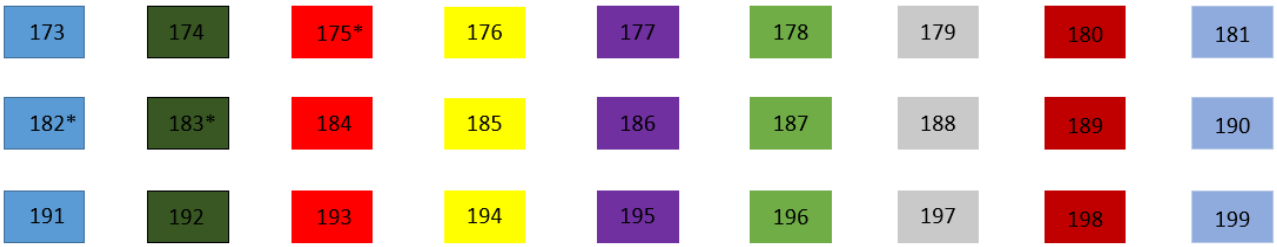

Figure 12: The figure exhibits the array of exons (most of these exons are only expressed in fetal titin) in the meta-only repeated region. The exons with identical sequences are marked with the same color and the numbers superscripted with asterisk signify exons that are inconsistently included in adult N2A isoform.

The pathophysiology mechanism of patient 2 is relatively difficult to characterize. Since the identified variant lies in meta-only exon, the disease onset was before birth but the muscle weakness persists throughout life without progression despite the fact that meta-transcript/fetal titin is replaced by adult skeletal titin isoform (Ottenheijm, Knottnerus et al. 2009).

If the transcripts generated from the allele harboring splicing mutation in intron 179 were subjected to degradation via NMD pathway due to splicing defect, the allele would be null, and the patient would not have survived until birth because titin kinase domain would be missing in homozygous state (the second variant c.100312_100315del p.(Lys33438Glyfs*4) creates PTC before titin kinase) which results in sarcomere disassembly (Gotthardt, Hammer et al. 2003).

Alternatively, if the mutation in intron 179 does not direct transcripts to NMD then underlying cause of disease could be the other splicing defect e.g. exon skipping or in-frame deletion, but this explanation is plausible for embryonic development as exons adjacent to intron 179 are only expressed in prenatal life and mutation in intron 179 is likely irrelevant for aberrant splicing in adult isoform. Nevertheless, meta-transcript only mutations have been postulated to have an impact on the structure and/or function of one or more unidentified titin developmental isoforms (Oates, Jones et al. 2018).

By the reason that skeletal muscles have a regenerative property (Carlson 1973), and the adult isoform is normal, the description of rudimentary is rather improbable in this particular case.

These explanations are mere speculations and no coherent description can be given without analyzing TTN isoforms from fetal muscles. Until now, the only clear genotype-phenotype correlation is that a mutation in meta-only exons causes specific early-onset (prenatal-onset) joint contractures called arthrogryposis (Fernandez-Marmiesse, Carrascosa-Romero et al. 2017), which is the attributed phenotype of this patient.

In patient 3, the first variant c.53354G>A p.(Trp17785*) in allele 1 generates PTC; hence this would be a null allele.

The production of slightly shorter protein (exon 67 skipped/loss of 93 amino acids) is expected from the second allele. It is to be noted that the higher number of reads in RNA-Seq data supporting the canonical exon 66-67 junction is supposedly from transcripts synthesized by the allele 1 that are degraded by NMD and their absence is contributing factor of disease phenotype. The splicing variant also causes a secondary effect i.e. skipping of exon 66 with 67, which is relatively a minor splicing event (supported by 544 reads, see table 7) but shows that splicing mutation may render its effect on more than a single exon in proximity. Both exons 66 and 67 are symmetric, and their exclusion does not alter the reading frame.

In Patient 4, the first variant c.35829_35830insA p.(Gly11944Argfs*7) of allele 1 in exon 164 generates PTC resulting from frameshift effect. However, the exon 164 is meta-only exon and would have directed the transcripts to undergo NMD only when included (mainly during prenatal

life). An attempt to elucidate the pathophysiology mechanism would be based mainly on conjectures as described for the case of patient 2.

The variant c.25063+1G>A in allele 2 causes intron 87 retention (8018 reads support intron retention, figure 11) and skipping of exons 86 and 87.

The skipping is quantitatively not a major event compared to reads supporting their inclusion (Table 7); however, the reads from other samples showed very low or no skipping of exons 86 and 87, which raises the question whether this skipping event is functionally relevant to the disease. The transcripts with intron 87 retained would be degraded via NMD because of PTC, whereas transcripts with exon(s) skipped would cause in-frame deletion of Ig-like domain(s) from the I-band region.

It is noteworthy that the higher number of reads from RNA-Seq data supporting exon 86-87 and 87-88 junctions (Table 7) is presumably from transcripts produced by allele 1, which are equivalent of wild type isoforms in adult life. We surmise that most transcripts from both alleles were targeted by NMD during embryonic life, though the patient survived because of exon 87 skipping event which intrinsically saved transcripts from degradation.

The splicing variant c.4646-1G>A in patient 5 prompts the loss of first 14 nucleotides of exon 27 by activating cryptic splice site. This out of frame effect p.(Ala1549Glyfs*21) will consequently trigger NMD and this would be a null allele.

The second variant c.107889delA p.(Lys35963Asnfs*9) also generates PTC as a consequence of frameshift, but in this particular case the transcripts are very unlikely to be targeted by NMD pathway as the premature stop codon appears in the last exon, as it has been hypothesized based on the results of previous studies that distal M-Band/C-terminal of titin harboring PTC does not undergo NMD degradation (Deo 2016).

Furthermore, if NMD targeted the transcripts generated from both alleles, the patient would not have survived until postnatal life. Thus, the underlying reason for the abnormal function of protein is a frameshift effect and a very small truncation, combined both would distort the M10 domain structure and are likely to obstruct the binding of proteins such as Obscurin, Obscurin-like 1 and Myospryn (Evila, Vihola et al. 2014).

Lastly, in patient 6 three distinct splicing effects were seen as a result of the first variant c.54190+1G>A in intron 281 (Table 8). Irrespective of their differences, all transcripts generated from the allele would be degraded by NMD as all three events cause disruption in the frame.

The second splicing variant c.107377+2dup causes retention of first nucleotide of intron 362 and induces PTC p.(Pro35793Argfs*4) but the transcripts are likely to evade NMD degradation due to the relative position of PTC (intron 362) and would result in production of stable truncated titin, as demonstrated by (Carmignac, Salih et al. 2007) that C-terminus titin deletion downstream of titin kinase domain (TK encoded by exon 359) still results in titin incorporation into sarcomere.

The current data on pathogenesis of titinopathies is scant, and data interpretation is often complicated. Mis-splicing causing premature stop and NMD is most likely pathogenic. The situation is more complex when a splicing variant results into a small in-frame deletion and a slightly shorter titin that is still near full-size protein. In such case, a mutation (small in-frame) could be pathogenic or well-tolerated.

6.4 GENOTYPE- PHENOTYPE CORRELATION

Since the genotype-phenotype correlation in above patients is precisely indicative of titinopathies; it does not rule out the possibility of other muscle genes acting as “genetic modifiers” for these phenotypes, and to which extent genetic modifiers (if any) may correlate with titinopathies phenotype. Titin itself acting as a genetic modifier in DCM has been proposed (Roncarati, Viviani Anselmi et al. 2013).

Even though the phenotypes of our patients are compatible with titinopathies, segregation studies are necessary for outright genotype-phenotype correlation (Savarese, Maggi et al. 2018), which we were unable to perform due to unavailability of family data and a small cohort of samples.

The explicit correlation, evident from our experimental data is that all patients have quasi-normal/near full-length (N2A) titin, demonstrating that a near full-length titin is vital for survival until birth. Furthermore, in patients 1-4 congenital or early childhood onset phenotypes were consistent with variants in A-band or I-band, whereas early adult/adult-onset phenotypes were observed when the variant is positioned in one of the last exons (patients 5 and 6). Furthermore, we reaffirm the already reported association between arthrogryposis and meta-only mutations.

Lastly, the disease onset with respect to the relative location of a variant in M-band of titin is not firmly established, and the variable ages of onset have been observed (manuscript in preparation).

6.5 LIMITATIONS

Considering the low prevalence of titinopathies (no official reports on disease prevalence available) and procuring muscle biopsies presents numerous obstacles, the number of samples analyzed in this project was fairly low. To overcome hindrances in obtaining muscle samples, we performed minigene splicing assay for the last 4 TTN exons (M-band) in HeLa cells. Minigene splicing assay is an experimental approach to transfect gDNA fragments and study splicing in mammalian cell lines.

Our plan was to observe constitutive splicing and in subsequent experiments perform site-directed mutagenesis to introduce mutations in canonical and cryptic splice sites to study the effects of potential splicing mutations. The splicing effects in TTN I-band studied via minigene splicing assay using HEK-293 and HeLa cell lines have been reported before (Ceyhan-Birsoy, Agrawal et al. 2013, Li, Guo et al. 2013). Nevertheless, we were unable to obtain any splicing results from HeLa cell line. As the genomic landscape of HeLa cell line is chaotic (Landry, Pyl et al. 2013), we speculate that HeLa cells don't express all splicing factors necessary for TTN M-band splicing. Upon literature search, we could not find any source reporting minigene splicing assay performed on TTN M-band region.

Our future intent is to select a cell line for minigene splicing assay that closely resembles human muscle cells such as C2C12 (mouse muscle progenitor cell line).

Another limitation of this study is that we could not attain results from variants located in meta-only repeated region, therefore their role in disease remains obscure, however, as clearly demonstrated in this thesis, the study of splicing variants affecting meta-only exons is not feasible using an adult muscle and probably requires a fetal tissue or use of a minigene splicing assay.

6.6 COMPARATIVE ANALYSIS OF DIFFERENT SEQUENCING METHODS TO STUDY TTN SPLICING VARIANTS

Five different sequencing methods were used in this project to characterize splicing variants. The pros and cons of methods are listed in table 9.

Among the traditional methods, direct Sanger sequencing was most convenient though many times the results were vague because of high background noise; thus, the PCR samples had to be gel extracted or cloned prior to sequencing. To a lesser extent, similar problem of low signal to noise ratio was encountered with gel extracted DNA sequencing. Nonetheless, PCR cloning, a much more complex and expensive procedure, always provided clear and unambiguous results.

Illumina RNA-Seq was chosen as a high-throughput sequencing approach as it has been reported as a superior technique over other transcriptomic approaches (Wang, Gerstein et al. 2009) and is apparently the most promising for TTN splicing variants analysis because of its high data throughput that enabled us to detect the highest number of splicing events, including minor splicing events and quantify them.

Regarding the efficacy of SMRT sequencing, one of the splicing effects (observed in RNA-Seq) was not seen in SMRT sequencing because of low reads count. This may be due to some technical aspects: a) pooling of several other PCR products that minimized the number of TTN reads in our results; b) we sequenced PCR fragments with a size range between 700 bp and 2000 bp, the smaller sized fragments were sequenced more often and generated higher number of reads than larger fragments. In our perspective, also considering the additional preliminary data not included in present thesis, SMRT can be very useful for routine diagnostic analysis of large individual transcripts such as those of dystrophin that nowadays are often sequenced via Sanger sequencing method using dozens of overlapping fragments.

Considering the intricacies of size and structure of TTN, An RT-PCR-based method to analyze the full-length transcript is not recommended. In our experience, whole transcriptome (Illumina RNA-Seq) is the most feasible and exhaustive approach to study TTN transcripts in patients with or without previously identified TTN variants.

Table 9: The table summarizes OUR EXPERIENCE with following sequencing techniques

	Direct Sanger sequencing	Gel DNA sequencing	Pcr cloning sequencing	Illumina RNA seq	SMRT
Strengths	High accuracy	High accuracy	High accuracy, clear results	High throughput and high accuracy, cost effective	Long reads (high resolution)
Drawbacks	Noisy data, ambiguous base calls	Noisy data, ambiguous base calls	High-cost	Short reads (low coverage), data analysis is tricky/mapping issues.	Single read is error prone, Low number of reads.

6.7 FUTURE PROSPECTS

Taking into account the immense size of TTN with 364 exons and the advancements in high-throughput sequencing technologies, it is logical to presume that the landscape of TTN splicing variants will greatly increase in the near future. However, the bigger hurdle in translation of current knowledge into clinical practice resides in variant classification rather than mere variant identification (Quintáns, Ordóñez-Ugalde et al. 2014). A study adjudged that 122 out of 460 mutations listed in literatures for severe childhood recessive diseases were common SNPs and sequencing errors (Bell, Dinwiddie et al. 2011); this underscores the significance of proper variant characterization and a more comprehensive approach for that (see conclusive remarks).

In addition, all patients analyzed in this project carried splicing variants in canonical splice sites; the higher uncertainty is associated with the characterization of splicing variants out of canonical splice sites and deep intronic/exonic variants mainly because variants in canonical splice sites are well known to disrupt normal splicing (Spurdle, Couch et al. 2008), and bioinformatics programs do not have high accuracy in predicting cryptic splice variants (Tang, Prosser et al. 2016), which as well drives attention towards the need for higher accuracy in silico prediction tools.

In a broader perspective, comprehensive studies are needed for proper evaluation of all structural domains of titin and their possible protein-protein interactions sites missed upon in-frame deletion, such data will provide deep insights into functional roles of individual titin domains and a vast understanding on pathogenesis of titinopathies.

7 CONCLUSIVE REMARKS

Based on guidelines and criteria established by American college of medical genetics and genomics for Mendelian disorders, a sequence variant can be classified as 1) pathogenic 2) likely pathogenic 3) variant of uncertain significance 4) likely benign 5) benign.

We used the online variant interpretation tool developed by (Kleinberger, Maloney et al. 2016), which is based on ACMG guidelines and classified the variants accordingly.

Table 10: The classification of TTN splicing variants studied

Variant	Clinical significance	Evidence
c.15776-1G>T (intron 54)	Pathogenic (Ia)	PVS1, PS3, PP4
c.67349-2A>C (intron 319)	Pathogenic (Ia) Likely pathogenic (II)	PVS1, PS3, PM3, PP4, PS3, PM3, PM4, PP4
c.37040-1G>A (intron 179)	VUS	
c.19426+2T>A(intron 67)	Likely pathogenic (II)	PS3, PM3, PM4, PP4
c.25063+1G>A (intron 87)	Pathogenic (Ia) Likely pathogenic (II)	PVS1, PS3, PP4 PS3, PM4, PP4
c.4646-1G>A (intron 26)	Pathogenic (Ia)	PVS1, PS3, PP4
c.54190+1G>A (intron 281)	Pathogenic (Ia)	PVS1, PS3, PP4
c.107377+2dup (intron 362)	Likely pathogenic (II)	PS3, PM3, PM4, PP4

The PVS1 criterion matched for the last variant (c.107377+2dup) but was replaced by PM4 because of the farther downstream position of the variant that is expected to cause small truncation without loss of function (LOF); ACMG also advises to exercise caution for such variants interpretation.

The variants rendering distinct splicing effects were classified as both pathogenic and likely pathogenic because of their disparate anticipated outcomes.

Finally, all variants classified as pathogenic include the PS3 criterion that requires functional studies of RNA or protein supportive of the damaging effect. However, functional studies merely at RNA and protein level have been questioned for variant classification of MYH7 gene, in the related panel discussion, the experts decided that strong functional evidence can only be established by analysis of a mammalian variant-specific knock-in model for MYH7 gene (Kelly, Caleshu et al. 2018).

Likewise, deeming the complexity of TTN, the interpretation of pathogenic effects is not straightforward as of other average-sized genes and it is difficult to infer pathogenicity in TTN solely on the basis of transcriptome studies and may require a variant-specific knock-in model to achieve 100% certainty.

As for which method is most suitable for TTN splicing variants characterization, our results illustrate that among the methods we used, no single method is ideal (though RNA-Seq is very efficacious) and it seems that combination of high-throughput and high-resolution (long reads) sequencing approaches is very propitious.

8 ACKNOWLEDGEMENTS

This master's thesis was conducted at Folkhalsän Research Center in a group studying neuromuscular disorders, led by Bjarne Udd.

I am very grateful to Bjarne Udd for giving me the opportunity to work in his group.

Special thanks to my immediate supervisor Marco Savarese and co-supervisor Per-harald Jonson for all the help during the entire duration of thesis work.

I also acknowledge the generous stipend provided by Samfundet Folkhalsän and Sigrid Jusélius Stiftelse for three and three consecutive months.

9 REFERENCES

Anna, A. and G. Monika (2018). "Splicing mutations in human genetic disorders: examples, detection, and confirmation." J Appl Genet 59(3): 253-268.

Ardui, S., A. Ameur, J. R. Vermeesch and M. S. Hestand (2018). "Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics." Nucleic acids research 46(5): 2159-2168.

Bang, M. L., T. Centner, F. Fornoff, A. J. Geach, M. Gotthardt, M. McNabb, C. C. Witt, D. Labeit, C. C. Gregorio, H. Granzier and S. Labeit (2001). "The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system." Circ Res 89(11): 1065-1072.

Bell, C. J., D. L. Dinwiddie, N. A. Miller, S. L. Hateley, E. E. Ganusova, J. Mudge, R. J. Langley, L. Zhang, C. C. Lee, F. D. Schilkey, V. Sheth, J. E. Woodward, H. E. Peckham, G. P. Schroth, R. W. Kim and S. F. Kingsmore (2011). "Carrier testing for severe childhood recessive diseases by next-generation sequencing." Science translational medicine 3(65): 65ra64-65ra64.

Boise, L. H., M. Gonzalez-Garcia, C. E. Postema, L. Ding, T. Lindsten, L. A. Turka, X. Mao, G. Nunez and C. B. Thompson (1993). "bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death." Cell 74(4): 597-608.

Burset, M., I. A. Seledtsov and V. V. Solovyev (2000). "Analysis of canonical and non-canonical splice sites in mammalian genomes." Nucleic acids research 28(21): 4364-4375.

Carlson, B. M. (1973). "The regeneration of skeletal muscle. A review." Am J Anat 137(2): 119-149.

Carmignac, V., M. A. Salih, S. Quijano-Roy, S. Marchand, M. M. Al Rayess, M. M. Mukhtar, J. A. Urtizberea, S. Labeit, P. Guicheney, F. Leturcq, M. Gautel, M. Fardeau, K. P. Campbell, I. Richard, B. Estournet and A. Ferreiro (2007). "C-terminal titin deletions cause a novel early-onset myopathy with fatal cardiomyopathy." Ann Neurol 61(4): 340-351.

Cazorla, O., A. Freiburg, M. Helmes, T. Centner, M. McNabb, Y. Wu, K. Trombitas, S. Labeit and H. Granzier (2000). "Differential expression of cardiac titin isoforms and modulation of cellular stiffness." Circ Res 86(1): 59-67.

Ceyhan-Birsoy, O., P. B. Agrawal, C. Hidalgo, K. Schmitz-Abe, E. T. DeChene, L. C. Swanson, R. Soemedi, N. Vasli, S. T. Iannaccone, P. B. Shieh, N. Shur, J. M. Dennison, M. W. Lawlor, J. Laporte, K. Markianos, W. G. Fairbrother, H. Granzier and A. H. Beggs (2013). "Recessive truncating titin gene, TTN, mutations presenting as centronuclear myopathy." Neurology 81(14): 1205-1214.

Chauveau, C., C. G. Bonnemann, C. Julien, A. L. Kho, H. Marks, B. Talim, P. Maury, M. C. Arne-Bes, E. Uro-Coste, A. Alexandrovich, A. Vihola, S. Schafer, B. Kaufmann, L. Medne, N. Hubner, A. R. Foley, M. Santi, B. Udd, H. Topaloglu, S. A. Moore, M. Gotthardt, M. E. Samuels, M. Gautel and A. Ferreiro (2014). "Recessive TTN truncating mutations define novel forms of core myopathy with heart disease." Hum Mol Genet 23(4): 980-991.

Chu, M., C. C. Gregorio and C. T. Pappas (2016). "Nebulin, a multi-functional giant." J Exp Biol 219(Pt 2): 146-152.

Deo, R. C. (2016). "Alternative Splicing, Internal Promoter, Nonsense-Mediated Decay, or All Three: Explaining the Distribution of Truncation Variants in Titin." Circ Cardiovasc Genet 9(5): 419-425.

Di Resta, C., S. Galbiati, P. Carrera and M. Ferrari (2018). "Next-generation sequencing approach for the diagnosis of human diseases: open challenges and new opportunities." EJIFCC 29(1): 4-14.

Evila, A., M. Arumilli, B. Udd and P. Hackman (2016). "Targeted next-generation sequencing assay for detection of mutations in primary myopathies." Neuromuscul Disord 26(1): 7-15.

Evila, A., A. Vihola, J. Sarparanta, O. Raheem, J. Palmio, S. Sandell, B. Eymard, I. Illa, R. Rojas-Garcia, K. Hankiewicz, L. Negrao, T. Lopponen, P. Nokelainen, M. Karppa, S. Penttila, M. Screen, T. Suominen, I. Richard, P. Hackman and B. Udd (2014). "Atypical phenotypes in titinopathies explained by second titin mutations." Ann Neurol 75(2): 230-240.

Fernandez-Marmiesse, A., M. C. Carrascosa-Romero, B. Alfaro Ponce, A. Nascimento, C. Ortez, N. Romero, L. Palacios, C. Jimenez-Mallebrera, C. Jou, S. Gouveia and M. L. Couce (2017). "Homozygous truncating mutation in prenatally expressed skeletal isoform of TTN gene results in arthrogryposis multiplex congenita and myopathy without cardiac involvement." Neuromuscul Disord 27(2): 188-192.

Freiburg, A. and M. Gautel (1996). "A molecular map of the interactions between titin and myosin-binding protein C. Implications for sarcomeric assembly in familial hypertrophic cardiomyopathy." Eur J Biochem 235(1-2): 317-323.

Gao, K., A. Masuda, T. Matsuura and K. Ohno (2008). "Human branch point consensus sequence is yUnAy." Nucleic acids research 36(7): 2257-2267.

Garvey, S. M., C. Rajan, A. P. Lerner, W. N. Frankel and G. A. Cox (2002). "The muscular dystrophy with myositis (mdm) mouse mutation disrupts a skeletal muscle-specific domain of titin." Genomics 79(2): 146-149.

Gautel, M. (2011). "Cytoskeletal protein kinases: titin and its relations in mechanosensing." Pflugers Arch 462(1): 119-134.

Ge, Y. and B. T. Porse (2014). "The functional consequences of intron retention: alternative splicing coupled to NMD as a regulator of gene expression." Bioessays 36(3): 236-243.

Gotthardt, M., R. E. Hammer, N. Hubner, J. Monti, C. C. Witt, M. McNabb, J. A. Richardson, H. Granzier, S. Labeit and J. Herz (2003). "Conditional expression of mutant M-line titins results in cardiomyopathy with altered sarcomere structure." J Biol Chem 278(8): 6059-6065.

Greaser, M. (2001). "Identification of new repeating motifs in titin." Proteins 43(2): 145-149.

Guo, W., S. J. Bharmal, K. Esbona and M. L. Greaser (2010). "Titin diversity--alternative splicing gone wild." J Biomed Biotechnol 2010: 753675.

Guo, W., S. Schafer, M. L. Greaser, M. H. Radke, M. Liss, T. Govindarajan, H. Maatz, H. Schulz, S. Li, A. M. Parrish, V. Dauksaite, P. Vakeel, S. Klaassen, B. Gerull, L. Thierfelder, V. Regitz-Zagrosek, T. A. Hacker, K. W. Saupe, G. W. Dec, P. T. Ellinor, C. A. MacRae, B. Spallek, R. Fischer, A. Perrot, C. Ozcelik, K. Saar, N. Hubner

and M. Gotthardt (2012). "RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing." Nat Med 18(5): 766-773.

Hackman, P., A. Vihola, H. Haravuori, S. Marchand, J. Sarparanta, J. De Seze, S. Labeit, C. Witt, L. Peltonen, I. Richard and B. Udd (2002). "Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin." American journal of human genetics 71(3): 492-500.

Hamdani, N., M. Herwig and W. A. Linke (2017). "Tampering with springs: phosphorylation of titin affecting the mechanical function of cardiomyocytes." Biophysical reviews 9(3): 225-237.

Herman, D. S., L. Lam, M. R. G. Taylor, L. Wang, P. Teekakirikul, D. Christodoulou, L. Conner, S. R. DePalma, B. McDonough, E. Sparks, D. L. Teodorescu, A. L. Cirino, N. R. Banner, D. J. Pennell, S. Graw, M. Merlo, A. Di Lenarda, G. Sinagra, J. M. Bos, M. J. Ackerman, R. N. Mitchell, C. E. Murry, N. K. Lakdawala, C. Y. Ho, P. J. R. Barton, S. A. Cook, L. Mestroni, J. G. Seidman and C. E. Seidman (2012). "Truncations of titin causing dilated cardiomyopathy." The New England journal of medicine 366(7): 619-628.

Keller, T. C., 3rd, K. Eilertsen, M. Higginbotham, S. Kazmierski, K. T. Kim and M. Velichkova (2000). "Role of titin in nonmuscle and smooth muscle cells." Adv Exp Med Biol 481: 265-277; discussion 278-281.

Kellermayer, D., J. E. Smith, 3rd and H. Granzier (2017). "Novex-3, the tiny titin of muscle." Biophysical reviews 9(3): 201-206.

Kelly, M. A., C. Caleshu, A. Morales, J. Buchan, Z. Wolf, S. M. Harrison, S. Cook, M. W. Dillon, J. Garcia, E. Haverfield, J. D. H. Jongbloed, D. Macaya, A. Manrai, K. Orland, G. Richard, K. Spoonamore, M. Thomas, K. Thomson, L. M. Vincent, R. Walsh, H. Watkins, N. Whiffin, J. Ingles, J. P. van Tintelen, C. Semsarian, J. S. Ware, R. Hershberger and B. Funke (2018). "Adaptation and validation of the ACMG/AMP variant classification framework for MYH7-associated inherited cardiomyopathies: recommendations by ClinGen's Inherited Cardiomyopathy Expert Panel." Genet Med 20(3): 351-359.

Keren, H., G. Lev-Maor and G. Ast (2010). "Alternative splicing and evolution: diversification, exon definition and function." Nat Rev Genet 11(5): 345-355.

Kleinberger, J., K. A. Maloney, T. I. Pollin and L. J. Jeng (2016). "An openly available online tool for implementing the ACMG/AMP standards and guidelines for the interpretation of sequence variants." Genet Med 18(11): 1165.

Krüger, M. and S. Kötter (2016). "Titin, a Central Mediator for Hypertrophic Signaling, Exercise-Induced Mechanosignaling and Skeletal Muscle Remodeling." Frontiers in physiology 7: 76-76.

Kulke, M., S. Fujita-Becker, E. Rostkova, C. Neagoe, D. Labeit, D. J. Manstein, M. Gautel and W. A. Linke (2001). "Interaction between PEVK-titin and actin filaments: origin of a viscous force component in cardiac myofibrils." Circ Res 89(10): 874-881.

Labeit, S., M. Gautel, A. Lakey and J. Trinick (1992). "Towards a molecular understanding of titin." Embo j 11(5): 1711-1716.

Labeit, S. and B. Kolmerer (1995). "Titins: giant proteins in charge of muscle ultrastructure and elasticity." Science 270(5234): 293-296.

Lahmers, S., Y. Wu, D. R. Call, S. Labeit and H. Granzier (2004). "Developmental control of titin isoform expression and passive stiffness in fetal and neonatal myocardium." Circ Res 94(4): 505-513.

Landry, J. J. M., P. T. Pyl, T. Rausch, T. Zichner, M. M. Tekkedil, A. M. Stütz, A. Jauch, R. S. Aiyar, G. Pau, N. Delhomme, J. Gagneur, J. O. Korbel, W. Huber and L. M. Steinmetz (2013). "The genomic and transcriptomic landscape of a HeLa cell line." G3 (Bethesda, Md.) 3(8): 1213-1224.

Lange, S., D. Auerbach, P. McLoughlin, E. Perriard, B. W. Schafer, J. C. Perriard and E. Ehler (2002). "Subcellular targeting of metabolic enzymes to titin in heart muscle may be mediated by DRAL/FHL-2." J Cell Sci 115(Pt 24): 4925-4936.

Lewartowski, B. and U. Mackiewicz (2014). "Cardiac titin, the giant sarcomeric protein in health and disease." Kardiol Pol 72(3): 215-222.

Li, H.-D., R. Menon, G. S. Omenn and Y. Guan (2014). "Revisiting the identification of canonical splice isoforms through integration of functional genomics and proteomics evidence." Proteomics 14(23-24): 2709-2718.

Li, S., W. Guo, C. N. Dewey and M. L. Greaser (2013). "Rbm20 regulates titin alternative splicing as a splicing repressor." Nucleic Acids Res 41(4): 2659-2672.

Lindeboom, R. G. H., F. Supek and B. Lehner (2016). "The rules and impact of nonsense-mediated mRNA decay in human cancers." Nature genetics 48(10): 1112-1118.

Linke, W. A. (2000). "Stretching molecular springs: elasticity of titin filaments in vertebrate striated muscle." Histol Histopathol 15(3): 799-811.

Linke, W. A. (2008). "Sense and stretchability: the role of titin and titin-associated proteins in myocardial stress-sensing and mechanical dysfunction." Cardiovasc Res 77(4): 637-648.

Linke, W. A. (2018). "Titin Gene and Protein Functions in Passive and Active Muscle." Annu Rev Physiol 80: 389-411.

Liu, Y., M. Gonzalez-Porta, S. Santos, A. Brazma, J. C. Marioni, R. Aebersold, A. R. Venkitaraman and V. O. Wickramasinghe (2017). "Impact of Alternative Splicing on the Human Proteome." Cell Rep 20(5): 1229-1241.

Ma, K. and K. Wang (2002). "Interaction of nebulin SH3 domain with titin PEVK and myopalladin: implications for the signaling and assembly role of titin and nebulin." FEBS Lett 532(3): 273-278.

Makarenko, I., C. A. Opitz, M. C. Leake, C. Neagoe, M. Kulke, J. K. Gwathmey, F. del Monte, R. J. Hajjar and W. A. Linke (2004). "Passive stiffness changes caused by upregulation of compliant titin isoforms in human dilated cardiomyopathy hearts." Circ Res 95(7): 708-716.

McElhinny, A. S., K. Kakinuma, H. Sorimachi, S. Labeit and C. C. Gregorio (2002). "Muscle-specific RING finger-1 interacts with titin to regulate sarcomeric M-line and thick filament structure and may have nuclear functions via its interaction with glucocorticoid modulatory element binding protein-1." The Journal of cell biology 157(1): 125-136.

Meyer, L. C. and N. T. Wright (2013). "Structure of giant muscle proteins." Frontiers in physiology 4: 368-368.

Miller, M. K., M. L. Bang, C. C. Witt, D. Labeit, C. Trombitas, K. Watanabe, H. Granzier, A. S. McElhinny, C. C. Gregorio and S. Labeit (2003). "The muscle ankyrin repeat proteins: CARP, ankrd2/Arpp and DARP as a family of titin filament-based stress response molecules." J Mol Biol 333(5): 951-964.

Musa, H., S. Meek, M. Gautel, D. Peddie, A. J. Smith and M. Peckham (2006). "Targeted homozygous deletion of M-band titin in cardiomyocytes prevents sarcomere formation." J Cell Sci 119(Pt 20): 4322-4331.

Nilsen, T. W. (2003). "The spliceosome: the most complex macromolecular machine in the cell?" Bioessays 25(12): 1147-1149.

Oates, E. C., K. J. Jones, S. Donkervoort, A. Charlton, S. Brammah, J. E. Smith, 3rd, J. S. Ware, K. S. Yau, L. C. Swanson, N. Whiffin, A. J. Peduto, A. Bournazos, L. B. Waddell, M. A. Farrar, H. A. Sampaio, H. L. Teoh, P. J. Lamont, D. Mowat, R. B. Fitzsimons, A. J. Corbett, M. M. Ryan, G. L. O'Grady, S. A. Sandaradura, R. Ghaoui, H. Joshi, J. L. Marshall, M. A. Nolan, S. Kaur, J. Punetha, A. Töpf, E. Harris, M. Bakshi, C. A. Genetti, M. Marttila, U. Werlauff, N. Streichenberger, A. Pestronk, I. Mazanti, J. R. Pinner, C. Vuillerot, C. Grosmann, A. Camacho, P. Mohassel, M. E. Leach, A. R. Foley, D. Bharucha-Goebel, J. Collins, A. M. Connolly, H. R. Gilbreath, S. T. Iannaccone, D. Castro, B. B. Cummings, R. I. Webster, L. Lazaro, J. Vissing, S. Coppens, N. Deconinck, H.-M. Luk, N. H. Thomas, N. C. Foulds, M. A. Illingworth, S. Ellard, C. A. McLean, R. Phadke, G. Ravenscroft, N. Witting, P. Hackman, I. Richard, S. T. Cooper, E.-J. Kamsteeg, E. P. Hoffman, K. Bushby, V. Straub, B. Udd, A. Ferreiro, K. N. North, N. F. Clarke, M. Lek, A. H. Beggs, C. G. Bönnemann, D. G. MacArthur, H. Granzier, M. R. Davis and N. G. Laing (2018). "Congenital Titinopathy: Comprehensive characterization and pathogenic insights." Annals of neurology 83(6): 1105-1124.

Ottenheijm, C. A. C., A. M. Knottnerus, D. Buck, X. Luo, K. Greer, A. Hoying, S. Labeit and H. Granzier (2009). "Tuning passive mechanics through differential splicing of titin during skeletal muscle development." Biophysical journal 97(8): 2277-2286.

Palmio, J., S. Leonard-Louis, S. Sacconi, M. Savarese, S. Penttilä, A.-L. Semmler, W. Kress, T. Mozaffar, T. Lai, T. Stojkovic, A. Berardo, R. Reisin, S. Attarian, A. Urtizberea, A. M. Cobo, L. Maggi, S. Kurbatov, S. Nikitin, J. C. Milisenda, F. Fatehi, M. Raimondi, F. Silveira, P. Hackman, K. G. Claeys and B. Udd (2019). "Expanding the importance of HMERF titinopathy: new mutations and clinical aspects." Journal of Neurology 266(3): 680-690.

Pan, Q., O. Shai, L. J. Lee, B. J. Frey and B. J. Blencowe (2008). "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing." Nat Genet 40(12): 1413-1415.

Park, E., Z. Pan, Z. Zhang, L. Lin and Y. Xing (2018). "The Expanding Landscape of Alternative Splicing Variation in Human Populations." Am J Hum Genet 102(1): 11-26.

Peled, Y., M. Gramlich, G. Yoskovitz, M. S. Feinberg, A. Afek, S. Polak-Charcon, E. Pras, B. A. Sela, E. Konen, O. Weissbrod, D. Geiger, P. M. Gordon, L. Thierfelder, D. Freimark, B. Gerull and M. Arad (2014). "Titin mutation in familial restrictive cardiomyopathy." Int J Cardiol 171(1): 24-30.

Peric, S., J. N. Glumac, A. Topf, D. Savic-Pavicevic, L. Phillips, K. Johnson, M. Cassop-Thompson, L. Xu, M. Bertoli, M. Lek, D. MacArthur, M. Brkusanin, S. Milenkovic, V. M. Rasic, B. Banko, R. Maksimovic, H. Lochmuller, V. R. Stojanovic and V. Straub (2017). "A novel recessive TTN founder variant is a common cause of distal myopathy in the Serbian population." Eur J Hum Genet 25(5): 572-581.

Plant, E. P., P. Wang, J. L. Jacobs and J. D. Dinman (2004). "A programmed -1 ribosomal frameshift signal can function as a cis-acting mRNA destabilizing element." Nucleic acids research 32(2): 784-790.

Puchner, E. M., A. Alexandrovich, A. L. Kho, U. Hensen, L. V. Schafer, B. Brandmeier, F. Grater, H. Grubmuller, H. E. Gaub and M. Gautel (2008). "Mechanoenzymatics of titin kinase." Proc Natl Acad Sci U S A 105(36): 13385-13390.

Quintáns, B., A. Ordóñez-Ugalde, P. Cacheiro, A. Carracedo and M. J. Sobrido (2014). "Medical genomics: The intricate path from genetic variant identification to clinical interpretation." Applied & translational genomics 3(3): 60-67.

Radke, M. H., J. Peng, Y. Wu, M. McNabb, O. L. Nelson, H. Granzier and M. Gotthardt (2007). "Targeted deletion of titin N2B region leads to diastolic dysfunction and cardiac atrophy." Proc Natl Acad Sci U S A 104(9): 3444-3449.

Raynaud, F., E. Fernandez, G. Coulis, L. Aubry, X. Vignon, N. Bleimling, M. Gautel, Y. Benyamin and A. Ouali (2005). "Calpain 1-titin interactions concentrate calpain 1 in the Z-band edges and in the N2-line region within the skeletal myofibril." Febs j 272(10): 2578-2590.

Roberts, A. M., J. S. Ware, D. S. Herman, S. Schafer, J. Baksi, A. G. Bick, R. J. Buchan, R. Walsh, S. John, S. Wilkinson, F. Mazzarotto, L. E. Felkin, S. Gong, J. A. MacArthur, F. Cunningham, J. Flannick, S. B. Gabriel, D. M. Altshuler, P. S. Macdonald, M. Heinig, A. M. Keogh, C. S. Hayward, N. R. Banner, D. J. Pennell, D. P. O'Regan, T. R. San, A. de Marvao, T. J. Dawes, A. Gulati, E. J. Birks, M. H. Yacoub, M. Radke, M. Gotthardt, J. G. Wilson, C. J. O'Donnell, S. K. Prasad, P. J. Barton, D. Fatkin, N. Hubner, J. G. Seidman, C. E. Seidman and S. A. Cook (2015). "Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease." Sci Transl Med 7(270): 270ra276.

Roncarati, R., C. Viviani Anselmi, P. Krawitz, G. Lattanzi, Y. von Kodolitsch, A. Perrot, E. di Pasquale, L. Papa, P. Portararo, M. Columbaro, A. Forni, G. Faggian, G. Condorelli and P. N. Robinson (2013). "Doubly heterozygous LMNA and TTN mutations revealed by exome sequencing in a severe form of dilated cardiomyopathy." European journal of human genetics : EJHG 21(10): 1105-1111.

Sakharkar, M. K., V. T. Chow and P. Kangueane (2004). "Distributions of exons and introns in the human genome." In Silico Biol 4(4): 387-393.

Satoh, M., M. Takahashi, T. Sakamoto, M. Hiroe, F. Marumo and A. Kimura (1999). "Structural analysis of the titin gene in hypertrophic cardiomyopathy: identification of a novel disease gene." Biochem Biophys Res Commun 262(2): 411-417.

Savarese, M., P. H. Jonson, S. Huovinen, L. Paulin, P. Auvinen, B. Udd and P. Hackman (2018). "The complexity of titin splicing pattern in human adult skeletal muscles." Skelet Muscle 8(1): 11.

Savarese, M., L. Maggi, A. Vihola, P. H. Jonson, G. Tasca, L. Ruggiero, L. Bello, F. Magri, T. Giugliano, A. Torella, A. Evila, G. Di Fruscio, O. Vanakker, S. Gibertini, L. Vercelli, A. Ruggieri, C. Antozzi, H. Luque, S. Janssens, M. B. Pasanisi, C. Fiorillo, M. Raimondi, M. Ergoli, L. Politano, C. Bruno, A. Rubegni, M. Pane, F. M. Santorelli, C. Minetti, C. Angelini, J. De Bleecker, M. Moggio, T. Mongini, G. P. Comi, L. Santoro, E. Mercuri, E. Pegoraro, M. Mora, P. Hackman, B. Udd and V. Nigro (2018). "Interpreting Genetic Variants in Titin in Patients With Muscle Disorders." JAMA Neurol 75(5): 557-565.

Savarese, M., J. Sarparanta, A. Vihola, B. Udd and P. Hackman (2016). "Increasing Role of Titin Mutations in Neuromuscular Disorders." J Neuromuscul Dis 3(3): 293-308.

Schafer, S., A. de Marvao, E. Adami, L. R. Fiedler, B. Ng, E. Khin, O. J. Rackham, S. van Heesch, C. J. Pua, M. Kui, R. Walsh, U. Tayal, S. K. Prasad, T. J. Dawes, N. S. Ko, D. Sim, L. L. Chan, C. W. Chin, F. Mazzarotto, P. J. Barton, F. Kreuchwig, D. P. de Kleijn, T. Totman, C. Biffi, N. Tee, D. Rueckert, V. Schneider, A. Faber, V. Regitz-Zagrosek, J. G. Seidman, C. E. Seidman, W. A. Linke, J. P. Kovalik, D. O'Regan, J. S. Ware, N. Hubner

and S. A. Cook (2017). "Titin-truncating variants affect heart function in disease cohorts and the general population." Nat Genet 49(1): 46-53.

Schrijver, I., N. Aziz, D. H. Farkas, M. Furtado, A. F. Gonzalez, T. C. Greiner, W. W. Grody, T. Hambuch, L. Kalman, J. A. Kant, R. D. Klein, D. G. Leonard, I. M. Lubin, R. Mao, N. Nagan, V. M. Pratt, M. E. Sobel, K. V. Voelkerding and J. S. Gibson (2012). "Opportunities and challenges associated with clinical diagnostic genome sequencing: a report of the Association for Molecular Pathology." J Mol Diagn 14(6): 525-540.

Spurdle, A. B., F. J. Couch, F. B. L. Hogervorst, P. Radice, O. M. Sinilnikova and I. U. G. V. W. Group (2008). "Prediction and assessment of splicing alterations: implications for clinical testing." Human mutation 29(11): 1304-1313.

Srebrow, A. and A. R. Kornblihtt (2006). "The connection between splicing and cancer." J Cell Sci 119(Pt 13): 2635-2641.

Tabish, A. M., V. Azzimato, A. Alexiadis, B. Buyandelger and R. Knoll (2017). "Genetic epidemiology of titin-truncating variants in the etiology of dilated cardiomyopathy." Biophys Rev 9(3): 207-223.

Tang, R., D. O. Prosser and D. R. Love (2016). "Evaluation of Bioinformatic Programmes for the Analysis of Variants within Splice Site Consensus Regions." Advances in bioinformatics 2016: 5614058-5614058.

Tanguay, R. L. and D. R. Gallie (1996). "Translational efficiency is regulated by the length of the 3' untranslated region." Mol Cell Biol 16(1): 146-156.

Tonino, P., B. Kiss, J. Strom, M. Methawasin, J. E. Smith, 3rd, J. Kolb, S. Labeit and H. Granzier (2017). "The giant protein titin regulates the length of the striated muscle thick filament." Nat Commun 8(1): 1041.

Toro, C., M. Olivé, M. C. Dalakas, K. Sivakumar, J. M. Bilbao, F. Tyndel, N. Vidal, E. Farrero, N. Sambuughin and L. G. Goldfarb (2013). "Exome sequencing identifies titin mutations causing hereditary myopathy with early respiratory failure (HMERF) in families of diverse ethnic origins." BMC neurology 13: 29-29.

Trombitas, K., M. Greaser, S. Labeit, J. P. Jin, M. Kellermayer, M. Helmes and H. Granzier (1998). "Titin extensibility in situ: entropic elasticity of permanently folded and permanently unfolded molecular segments." J Cell Biol 140(4): 853-859.

Tskhovrebova, L. and J. Trinick (2003). "Titin: properties and family relationships." Nat Rev Mol Cell Biol 4(9): 679-689.

Tskhovrebova, L. and J. Trinick (2010). "Roles of titin in the structure and elasticity of the sarcomere." J Biomed Biotechnol 2010: 612482.

Udd, B., A. Vihola, J. Sarparanta, I. Richard and P. Hackman (2005). "Titinopathies and extension of the M-line mutation phenotype beyond distal myopathy and LGMD2J." Neurology 64(4): 636-642.

Wang, E. T., R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth and C. B. Burge (2008). "Alternative isoform regulation in human tissue transcriptomes." Nature 456(7221): 470-476.

Wang, Z. and C. B. Burge (2008). "Splicing regulation: from a parts list of regulatory elements to an integrated splicing code." RNA (New York, N.Y.) 14(5): 802-813.

Wang, Z., M. Gerstein and M. Snyder (2009). "RNA-Seq: a revolutionary tool for transcriptomics." Nat Rev Genet 10(1): 57-63.

Young, P., C. Ferguson, S. Bañuelos and M. Gautel (1998). "Molecular structure of the sarcomeric Z-disk: two types of titin interactions lead to an asymmetrical sorting of alpha-actinin." The EMBO journal 17(6): 1614-1624.

10 ELECTRONIC REFERENCES

http://www.musclegenetable.fr/index.html

https://www.cardiodb.org/titin/

http://primer3.ut.ee/

https://www.medschool.umaryland.edu/Genetic_Variant_Interpretation_Tool1.html/

11 APPENDICES

PRIMERS SEQUENCES	TARGETED VARIANTS
25F GGTAGACCTATGCCAGAGACGTTC	c.4646-1G>A
28R CCTCCAGTCACGCATTCTTACTT	
54F AGCTTTTCCAATGGTGTTGCAGT	c.15776-1G>T
56R CGACAAGTGAAATTCCCTGCATCA	
57R AGAACATCCTTTGAGCCGGGTT	
64F TGCCAGGGTAGAAAATAGTGG	c.19426+2T>A
69R GTCGCCCAGGTTTCACTAGGAA	
85-86F (Junction primer) GTGAAAGAACCAGCCACCTTTGT	c.25063+1G>A
89-90R (Junction primer) GTTCTTGTACACCCAGCTGAGC	
170F (Repetitive region) GAGATTATCCCTGAAAAGAAAGTGTCG	c.37040-1G>A
182R (Repetitive region) CTTTTGTGGGTGGCACTTCAGG	
280F TCACACTTGACTGGAAAGAGCCC	c.54190+1G>A
283R AGATTTCTTGGTGGGGATGGGC	
319F ACTTCTTCCGAGTGTTTGCTGA	c.67349-2A>C
321R TTATGGGGATCTTAAGCCGGAAGTT	
361F GATGCTCCAGCCTTCATCTC	c.107377+2dup
364R (CATGTTACTTCTGGGGTAGGC)	

GENES ANALYZED IN MYOCAP ASSAY

ABHD5
ACADS
ACADVL
ACTA1
ACTN3
ACVR1
AGL
AIFM1
ALG13
ALG14
ALG2
ANO5
ATP2A1
B3GALNT2
DCST24
DCST23
DCST22
DCST21
DCST20
DCST19
DCST18
DCST17
DCST16
DCST15
DCST14
DCST13
DCST12
DCST11
DCST10
DCST9
DCST8
DCST7
DCST6
DCST5
DCST4
DCST3
DCST2
DCST1
DCST0
DCST1
DCST2
DES
DMD
DMPK
DNAJB6
DNM2
DOLK
DPAGT1
DPM1
DPM2
DPM3
DYSF
ECEL1
EMD
ENO3
ETFA
ETFB
ETFDH
EXOSC3
FBN2
FBXO32
FHL1
FHL2
FKRP
FKTN
FLNC
GAA
GATM
GBE1
GMPPB
GNB4
GNE
GOSR2
GSN
GTDC2 (POMGNT2)
GYS1
HEXB
HK1
HSPB8
ISCU
ISPD
ITGA7
ITPR1
KBTBD10 (KLHL41)
KBTBD13
KBTBD5 (KLHL40)
KCNE1
KCNE3
KCNJ2
KCNQ1
KLHL9
KY
LAMA2
LAMB1
LAMP1
LAMP2
LARGE
LDB3
LDHA
LDHB
LIFR
LIMS2
LMNA
MARS
MATR3
MBNL1
MBNL2
MBNL3
MEGF10
MSTN
MTM1
MTMR14
MYBPC1
MYBPC2
MYBPC3
MYH1
MYH2
MYH3
MYH4
MYH7
MYH8
MYL1
MYL10
MYL12A
MYL12B
MYL2
MYL3
MYL4
MYL5
MYL6
MYL6B
MYL7
MYL9
MYLIP
MYLK
MYLK2
MYLK3
MYLK4
MYLPF
MYOM1
MYOM2
MYOM3
MYOT
MYOZ1
MYOZ2
MYOZ3
MYPN
NBR1
NEB
NEBL
NEFL
NTRK1
OBSCN
OBSL1
PABPN1
PDK3
PDLIM3
PDLIM5
PDLIM7
PFKM
PGAM2
PGK1
PGM1
PLEC

PLEKHG4
PLEKHG5
PLN
PNPLA2
POLG
POMGNT1
POMT1
POMT2
PRKAG2
PTPLA
PTRF
PYGM
QDPR
RBCK1
RYR1
SBF1
SCN4A
SEPN1
SGCA
SGCB
SGCD
SGCG
SGK196 (POMK)
SLC22A5
SLC25A20
SLC5A2
SMCHD1
SOX10
SQSTM1
SRF
STAC3
STIM1
STIM2
SYNE1
SYNE2
SYNE3
SYNPO2
TCAP
TFG
TIA1
TIAL1
TMEM43
TMEM5
TMEM55A
TMEM55B
TMOD3
TNNC1
TNNC2
TNNI1
TNNI2
TNNI3
TNNT1
TNNT3
TNPO3
TNXB
TOR1AIP1
TPM1
TPM2
TPM3
TPP1
TRAPPC1
TRAPPC11
TRIM32
TRIM54
TRIM55
TRIM63
TTN
TTR
UBA1
VCP
VMA21
VRK1

cDNAs POOLED WITH TTN SAMPLE IN SMRT RUN

ACTN2
ANO5A
ANO5B
CAPN3
DES
DYSF
ENO3
FLNC
GNE
HSPG
ITGA7
LDB3
MYH7
POPDC2
SACS
SGCG
SQSTM1
STAC3
TIA1
TNNT1
TPM2
TPM3
VCP
CRYAB
FXR1
MATR
MBLN1
MYPN
MACF1
SGCB
SMPX